UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ROLE OF CAUDAL-RELATED HOMEOBOX (CDX) TRANSCRIPTION FACTORS IN TRUNK NEURAL CREST DEVELOPMENT

DISSERTATION

PRESENTED

AS PARTIAL REQUIREMENT

OF THE DOCTORATE OF BIOCHEMISTRY

BY

ORALY SANCHEZ-FERRAS

AUGUST 2015

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ANALYSE DES FONCTIONS NEURALES DES PROTÉINES CDX

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN BIOCHIMIE

PAR

ORALY SANCHEZ-FERRAS

AOÛT 2015

To my family and friends, specially to Alejandro

ACKNOWLEDGEMENTS

This doctoral thesis is the result of the work of a large network of people, involving the combined input of professors, labmates, family and friends. First, I would like to thank my director Dr. Nicolas Pilon for his guidance, support and patience, for the great lessons and discussions about embryogenesis, for the career advice, for helping me grow and for showing me that making science is hard work and requires devotion, humility, integrity and perseverance. I would also like to thank Dr. Benoît Barbeau, Dr. Joanne Paquin and Dr. David Lohnes for joining the evaluation committee and revising the manuscript. A special thank you to my labmates for their support and patience, in particular to those who worked very hard to make this project progress: Baptiste Coutaud, Taraneh Djavanbakht Samani, Isabelle Tremblay, Ouliana Souchkova, Guillaume Bernas, Omar Farnos, Emilie Laberge-Perrault, Aboubacrine M. Touré, Dr. Mélanie Béland. Support and advice from other members of the BioMed research center— especially Denis Flipo, professors Louise Brissette and François Dragon— and from Dr. Maxime Bouchard (McGill University) was also essential.

Finally, a very special thank you to my extended family and friends — my Cuban family and friend network, my friends Mari, Ruben, Aida, Yoisel, Miry, Suzanne, Marianne, and professors throughout my student life —, for helping and encouraging me to get here. To my dear love and best friend Alejandro: words are not enough to thank you for your support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOTATIONS

aa	Amino acid
AAA	Anterior arch of the atlas
Alk3	Activin-like kinase 3
AME	Anterior mesendoderm
ANP	Anterior neural plate
A-P	Anterior-Posterior
APC	Adenomatosis polyposis coli
AVE	Anterior visceral endoderm
bFGF	Basic fibroblast growth factor
bHLH	Basic helix-loop-helix
BMP	Bone Morphogenetic Protein
BMPR1	BMP receptor type I
BMPR2	BMP receptor type II
Cad	*Caudal* gene family
CAKUT	Congenital anomalies of the kidney and urinary tract
cAMP	Cyclic adenosine monophosphate
CBP	cAMP-response-element binding protein (CREB)-binding protein
cDNA	Complementary Deoxyribonucleic acid
Cdx	Caudal-related homeobox
CdxBS	Cdx Binding Site
CE	Convergent-extension
Cer1	Cerberus-like protein 1
ChIP	Chromatin immunoprecipitation
CHX	Cycloheximide
Ci	Cubitus interruptus
CK1	Casein kinase-1
cKO	Conditional knockout
CNH	Chordoneural hinge
CNE	Conserved non-coding element
CNS	Central nervous system
CRM	Cis-regulatory module
Cos2	Costal 2

COUP-TFs	Chicken ovalbumin upstream promoter-transcription factors
CRMs	cis-regulatory modules
Dct	Dopachrome tautomerase
DE	Distal enhancer
Dhh	Desert Hedhehog
Dkk1	Dickkopf 1
DLHPs	Dorsolateral hinge points
DNA	Deoxyribonucleic acid
Dpp	Decapentaplegic growth factor
DRG	Dorsal root ganglia
Dsh/Dvl	Dishevelled
D-V	Dorsal-Ventral
DVE	Distal visceral endoderm
e	Embryonic day
ECR	Evolutionary conserved region
Edn3	Endothelin 3
Ednrb	Endothelin receptor type b
EGFP	Enhanced green fluorescent protein
EGO	Early gastrula organizer
EMT	Epithelial to mesenchymal transition
EMSA	Electrophoretic mobility shift assay
ES	Embryonic stem cell
EnR	Engrailed
EnRCdx1	Engrailed-Cdx1 homeodomain fusion protein
ERBB3	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3
ExE	Extraembryonic ectoderm
FACS	Fluorescence-activated cell sorting
FGF	Fibroblast growth factor
FLRG	Follistatin-related gene
FoxD3	Forkhead Box D3
FP	Floor plate
FRPs	Frizzled-related proteins
Fz	Frizzled
GFP	Green fluorescent protein
Gli	Glioma-associated oncogen

Gli1	Glioma-associated oncogen homolog 1
Gli2	Glioma-associated oncogen homolog 2
Gli3	Glioma-associated oncogen homolog 3
GliA	Gli activator
GliR	Gli repressor
GO	Gastrula organizer
GPCR	G protein-coupled receptor
GSK3	Glycogen synthase kinase-3
GST	Glutathione S -transferase
Hcn3	Hyperpolarisation-activated cation-3 channel
H&E	Hematoxylin and eosin
Hh	Hedhehog
HIPK2	Homeodomain-interacting protein kinase 2
HMG	High mobility group
ICC-LCs	Interstitial cells of Cajal-like cells
ICM	Inner cell mass
Ihh	Indian Hedhehog
IRES	Internal ribosome entry site
kDa	kiloDalton
kb	kilobase
KE	Cardiac enhancer
KO	Knockout
LDA	Ligand-dependent antagonism
LEF/TCF	Lymphoid enhanced factor/ T-cell factor
Lefty 1	left-right determination factor 1
L-R	Left-Right
LRE	Lef/Tcf response element
LRP5/6	Low-density-lipoprotein-related protein 5/6
MAPK	Mitogen-activated protein kinase
mES	mouse embryonic stem cells
Mgf	mast cell growth factor
MGO	Midgastrula Organizer
MHP	Median hinge point
Mitf	Microphtalmia associated transcription factor
mRNA	messenger Ribonucleic acid

MSA	Migration Staging Area
MSCs	melanocyte stem cells
Msh/Msx	Muscle segment homeobox
NC	Neural crest
N-CAM	Neural-cell adhesion molecule
NCE	Neural crest enhancer
NC-GRN	Neural crest-gene regulatory network
N2a	Neuro 2a
NF-kB	Nuclear factor kappa Beta
NT	Neural tube
NTDs	Neural tube defects
opa	odd-paired
Pax	Paired box transcription factor
Pax3	paired box 3 transcription factor
Pax7	paired box 7 transcription factor
PCR	Polymerase chain reaction
PCP	Planar cell polarity
PE	Proximal enhancer
PGC	primordial germ cells
PNP	Posterior neural plate
PrE	Primitive endoderm
PS	Primitive streak
PSM	Presomitic mesoderm
pSmad1/5/8	phosphorylated-Smad1/5/8
Ptch1	Patched 1
R2	rhombomere 2
RA	Retinoic acid
RARE	RA response element
RNA	Ribonucleic acid
ROR1/2	Receptor tyrosine kinase-like orphan receptor
R-Smad	Receptor activated Smad
RT-PCR	Reverse transcription- polymerase chain reaction
RXR	Retinoid X receptor
RYK	Receptor-like tyrosine kinase
7TM	Seven-transmembrane
Shh	Sonic hedgehog

SMC	Smooth muscle cell
Smo	Smoothened
Smad	Fusion of names Sma (from small body size protein)-Mad (from mothers against decapentaplegic protein)
Sox	SRY (Sex determining region Y)-box
Sp	Splotch
Sp1	Specificity protein 1
ß-TrCP	beta-tranducin repeats containing protein
T	Brachyury
TBP	TATA box binding protein
TE	Trophectoderm
TGF-ß	Transforming Growth Factor-beta
Trp-1	Tyrosine-related protein-1
TSS	Transcription start site
Tyr	Tyrosinase
VE	visceral endoderm
WIF	Wnt-inhibitor protein
Wnt	Wingless/int1-related
WS	Waardenburg syndrome
ZF	Zinc finger
ZF-NC	Zinc finger N-terminally conserved domain
Zic	Zinc finger of the cerebellum
ZicBS	Zic binding site
ZOC	Zic/Odd paired conserved motif

RÉSUMÉ

Chez les vertébrés, une première ébauche du système nerveux central est obtenue par la neurulation qui mène à la formation du tube neural (TN) à partir de la plaque neurale (ou neuroectoderme). La neurulation est couplée à la spécification de l'identité positionnelle des cellules le long de l'axe antérieur-postérieur ainsi qu'au développement des cellules de la crête neurale (CCN) aux extrémités latérales de la plaque neurale. Les CCN migrent de la partie dorsale du TN et colonisent différentes régions de l'embryon. Ces cellules multipotentielles forment, entre autres, la totalité du système nerveux périphérique ainsi que les mélanocytes. Le développement des CCN est contrôlé par un réseau complexe de voies de signalisation impliquant l'intégration des mêmes signaux extracellulaires qui contrôlent l'établissement de l'axe antérieur-postérieur (ex : la voie Wnt canonique) et l'axe dorsal-ventral (ex : la voie BMP). Il est connu que ces signaux induisent l'expression d'un groupe important de facteurs de transcription, notamment Pax3/7, Msx1/2 et Zic2 lesquels ensuite activent l'expression des spécificateurs de la crête neurale tels que Sox9/10 et FoxD3. Toutefois, la façon dont ce réseau moléculaire est intégré au niveau transcriptionnel dans la région du tronc de l'embryon n'est pas bien comprise.

Les gènes *Cdx* (*Cdx1*, *Cdx2* et *Cdx4*) encodent des facteurs de transcription à homéodomaine fortement exprimés dans le neuroectoderme pendant l'ontogenèse des CCN du tronc. Toutefois, les fonctions neurales des gènes *Cdx* chez la souris sont largement méconnues à cause de leur redondance fonctionnelle et de la létalité embryonnaire précoce des doubles et triples mutants *Cdx*-nuls. Durant le développement de l'axe antérieur-postérieur, les facteurs de transcription Cdx intègrent plusieurs signaux postériorisants, plus particulièrement ceux de la voie Wnt canonique, afin de les relayer sur le promoteur des gènes *Hox* pour que ces derniers spécifient les différents segments de l'embryon. Dans cette thèse de doctorat, nous avons émis l'hypothèse que les facteurs Cdx sont d'importants régulateurs du développement des CCN chez la souris et occupent une position stratégique dans le réseau moléculaire en tant qu'intégrateurs des signaux

inducteurs. Dans une première étude, nous démontrons que les protéines Cdx agissent comme intermédiaires dans le mécanisme de régulation de l'expression neurale de *Pax3* par la voie Wnt canonique au niveau du tronc et régulent directement l'expression de *Pax3* en se liant sur une courte séquence d'ADN régulatrice située dans le promoteur proximal de *Pax3*, appelée *Neural Crest Enhancer 2* (NCE2). Dans une deuxième étude nous avons dévoilé le réseau moléculaire qui agit sur NCE2 et qui fait en sorte que *Pax3* s'exprime au bon endroit et au bon moment dans le neuroectoderme. Nous démontrons que le facteur de transcription Zic2 est un nouveau régulateur direct de *Pax3* via NCE2, que les protéines Cdx agissent en partenariat avec les facteurs de transcription Sox2 et Zic2 et que ce denier agit en fait sous le contrôle de la voie Shh. Cette étude est la première à apporter les bases moléculaires de l'intégration de signaux de l'axe antérieur-postérieur (Wnt) et de l'axe dorsal-ventral (Shh) sur un seul module régulateur de *Pax3*. Afin de dévoiler un rôle direct des protéines Cdx dans le développement des CCN *in vivo*, nous avons généré au cours d'une troisième étude un nouveau modèle de souris (R26R-$_{FLAG}$EnRCdx1) exprimant un dominant négatif Cdx (EnRCdx1) dans le TN dorsal, incluant les CCN. Cette approche nous a permis de dévoiler un rôle des protéines Cdx dans le contrôle du développement des mélanocytes ainsi que d'autres populations dérivées des CCN tels que les ganglions myentériques. Nos analyses suggèrent que les protéines Cdx agissent tôt dans les premières étapes du réseau moléculaire qui contrôle le développement des CCN en régulant directement l'expression de *Pax3*, ainsi que d'autres nouvelles cibles (*Msx1* et *FoxD3*) via leur promoteur proximal. Somme toute, cette étude démontre pour la première fois un rôle important et direct des protéines Cdx dans le développent des CCN du tronc.

Mots clés : Voie Wnt canonique, Cdx, EnRCdx1, crête neurale, réseau moléculaire, NCE2, Pax3, Zic2, Msx1, FoxD3, souris.

ABSTRACT

During vertebrate embryogenesis, the nervous system develops from the neuroectoderm by coordination of several processes following neural induction. Concomitantly with embryo elongation by addition of new cells from the posterior growth zone, the neural plate begins to roll and form the neural tube (NT), the precursor of the central nervous system (CNS) whereas at the lateral borders of the neural plate is induced an amazing population of multipotent and highly migratory cells known as the neural crest (NC). After closure of the NT, newly formed NC cells (NCC) localize at its dorsal part from where they delaminate and emigrate in order to colonize the embryo. Also known as the fourth germ layer, NCC give rise to multiple structures and mature cell types, including, to name a few, glia and neurons of the peripheral and enteric nervous system as well as melanocytes of the skin. NC development is controlled by a complex gene regulatory network (GRN) involving the combinatorial input of Anterior-Posterior (e.g. canonical Wnt pathway) and Dorsal-Ventral (e.g. BMP pathway) positional cues that lead to the stepwise induction of sets of genes encoding transcription factors known as the neural plate border specifiers (e.g. Pax3/7, Msx1/2 and Zic2) and the NC specifiers (e.g. Sox9/10 and FoxD3). However, how the NC-GRN is orchestrated at the transcriptional level in the trunk region of the vertebrate embryo is not well understood.

In this doctoral thesis, we tested the hypothesis that the *Caudal-related homeobox (Cdx)* genes (*Cdx1*, *2* and *4*) are important regulators of trunk NC development in the mouse and act at the top of the trunk NC-GRN to convey positional information from the signaling module to the promoter of NC-regulatory genes. In a first study, we demonstrated that Cdx act downstream of canonical Wnt signaling to directly induce *Pax3* expression in the caudal neuroectoderm via binding to a short trunk NC enhancer of the *Pax3* proximal promoter called NCE2. In a second study, we extended this knowledge by showing that the zinc finger transcription factor Zic2 is

also a direct regulator of *Pax3* NCE2 and a Cdx neural cofactor. This work led us to propose that a regulatory circuit involving the integration of a posteriorizing input from the canonical Wnt-Cdx pathway, a Dorso-Ventral instructive cue from the Shh-Nkx6.1-Zic2 pathway and a general neural input from the SoxB member Sox2 is involved in the establishment and refinement of *Pax3* expression domain in the caudal neuroectoderm. To demonstrate a direct Cdx role in the control of NC development *in vivo*, we generated a novel conditional pan-Cdx loss-of-function mouse line (R26R-$_{FLAG}$EnRCdx1) expressing a Cdx repressor fusion protein (EnRCdx1) under the control of the *ROSA26* promoter and targeted expression of EnRCdx1 to the NC tissue. This approach efficiently circumvented Cdx functional overlap and the early embryonic lethality of *Cdx* compound mutants, showing for the first time that Cdx proteins directly impact on NC development in the mouse, notably on development of skin melanocytes and myenteric neural ganglia. Mechanistic studies indicate that Cdx are at the head of the NC-GRN and directly control expression of two novel targets *Msx1* and *FoxD3* also via proximal promoter regulatory regions.

In summary, this work uncovers a direct role for Cdx transcription factors in trunk NC development.

Key words: Canonical Wnt, Cdx, EnRCdx1, neural crest, neural crest enhancer, gene regulatory network, Zic2, Pax3, Msx1, FoxD3, mouse.

INTRODUCTION

0.1 Three major signaling pathways in embryo development

0.1.1 Wnt signaling

As will be evidenced throughout this work, Wnt signaling plays essential roles in embryonic development in the control of body axis patterning, cell-fate specification, proliferation, differentiation, cell polarity, morphogenesis and stem cell renewal (for reviews see (Clevers and Nusse, 2012; Hikasa and Sokol, 2013)). In adulthood, Wnt signaling is involved in the regulation of tissue homeostasis. Embryogenesis and life cannot normally go on when Wnt signaling is deregulated. Mutation in members of the Wnt pathways causes catastrophic developmental defects that lead to congenital anomalies such as spina bifida and in adult life serious diseases such as cancer (reviewed in Clevers and Nusse, 2012).

Wnt signaling has been evolutionarily conserved from the origin of metazoans (Clevers and Nusse, 2012). Since unicellular organisms lack *Wnt* genes, it is considered that emergence of Wnt signaling was an important event in the origin of multicellular animals (Petersen and Reddien, 2009). The term Wnt is the result of the fusion of the name of the *Drosophila* segment polarity gene *wingless* and the name of the vertebrate homolog *integrated or int-1* (Nusse and Varmus, 1982; Rijsewijk et al., 1987). *Int-1* (further called *Wnt1*), was the first vertebrate *Wnt* gene to be identified in virally induced breast tumors (Nusse and Varmus, 1982), followed by the discovery of Drosophila *wingless* (Rijsewijk et al., 1987). From that date on, 19 *Wnt* genes have been identified in mammalian genomes, classified in 12 subfamilies according to similarity in amino acid sequence (Clevers and Nusse, 2012; Gordon and Nusse, 2006). Cnidarians contain at least 11 of these subfamilies, thus demonstrating the high degree of conservation of Wnt signaling all over the animal kingdom (Gordon and Nusse, 2006).

Wnt ligands are short range, cysteine-rich, secreted glycoproteins (~40 kDa in size) that need to be glycosylated and palmitoylated prior to release to the extracellular milieu (Takada et al., 2006; Tanaka et al., 2002). It has been shown that the highly conserved protein Porcupine mediates palmitoylation of Wnt proteins in the endoplasmic reticulum of the producing cell, whereas secretion to the extracellular milieu is mediated by the seven-transmembrane Wntless protein and the Retromer complex (reviewed in Clevers and Nusse, 2012; Komiya and Habas, 2008).

Once secreted, Wnt ligands interact with target cells via binding to a heterodimeric receptor complex formed by a Frizzled (Fz) and the low-density-lipoprotein-related protein 5/6 (LRP5/6) protein. Particularly, it has been shown that Wnt ligands bind to the large extracellular N-terminal cysteine-rich domain of the Fz receptor family, which is composed of ten members in humans and are seven-pass transmembrane (7TM) receptors. Vertebrate LRP5/6 is a single-pass transmembrane protein and recent studies suggest that LRP5/6 may also contain binding sites for Wnt ligands (Gong et al., 2010). Binding of Wnt to the receptor complex is an essential step in Wnt signaling transduction and a key point of regulation. Several extracellular antagonists of Wnt signaling act at this level to exclude Wnt signaling from certain regions or to fine-tune signaling thresholds. These include the soluble Frizzled-related proteins (sFRPs), Wnt-inhibitor protein (WIF) – both of which bind Wnts – or the protein Dikkopf (Dkk) that binds to LRP5/6 (Bovolenta et al., 2008; Glinka et al., 1998).

Binding of Wnt to the receptor complex recruits the cytoplasmic phosphoprotein Dishevelled (Dsh/Dvl); which has been shown to directly interact with the cytoplasmic domain of Fz (Chen et al., 2003). At this point, Wnt signaling transduction branches into three major pathways: the canonical or Wnt/ß-catenin pathway, the Planar Cell Polarity pathway and the Wnt-Ca^{2+} pathway (reviewed in Komiya and Habas, 2008). Although Dsh is a major component of the three branches,

it is still not well understood how this enigmatic protein regulates and makes the switch between the three pathways (Komiya and Habas, 2008).

Wnt1, Wnt3 and Wnt8 are recognized as examples of canonical ligands that bind to the Fz-LRP heterodimer and activate the ß-catenin pathway (Hikasa et al., 2010; MacDonald et al., 2009). Other Wnt proteins, such as Wnt4, Wnt5 and Wnt11 are recognized as examples of non-canonical members and have been established to trigger the PCP pathway (Wu and Mlodzik, 2009). Non-canonical ligands recruit other receptors than Fz, including members of the receptor tyrosine kinase family ROR1/2 and RYK (reviewed in Gordon and Nusse, 2006; Hikasa and Sokol, 2013). However, other studies indicate that Wnt11 can also signal via the canonical branch (Tao et al., 2005). These and other observations thus suggests that activation of one Wnt branch or the other depends on the cell type and the receptors expressed in a given cell rather than the nature of the Wnt ligand (Grumolato et al., 2010; van Amerongen et al., 2008). Of the Wnt branches, the canonical has been the most extensively studied and the one involved in specification of the body axis (Hikasa and Sokol, 2013). Because the Wnt/ß-catenin pathway is the most relevant to this work, a detailed description of this pathway follows, as well as a general description of the non canonical PCP arm.

0.1.1.1 Canonical Wnt/ß-catenin pathway

The hallmark of the canonical Wnt pathway is the accumulation of the protein ß-catenin in the cytoplasm and translocation into the nucleus (Clevers and Nusse, 2012). It is necessary to mention that ß-catenin is also a major component of cadherin-based adherens junctions. In this regard, it has been demonstrated that the pool of cadherin-bound ß-catenin can also be used for canonical Wnt signaling (Kam et al., 2009). In the Wnt-canonical arm, ß-catenin is constantly produced, but its stability and thus the output of the signal is regulated by a cytoplasmic destruction complex. The core components of the destruction complex are the tumor suppressor

proteins Axin and adenomatosis polyposis coli (APC) as well as two constitutively active serine-threonine kinases: glycogen synthase kinase-3 (GSK3) and casein kinase-1 (CK1) (Clevers and Nusse, 2012). Axin is the central scaffold of the destruction complex interacting with all other core components (Lee et al., 2003). APC is a large multi-domain protein that also has scaffolding properties and interacts with both ß-catenin and Axin (Lee et al., 2003). Mutations in AXIN, APC and ß-catenin also result in cancer (reviewed in Clevers and Nusse, 2012). Particularly, mutations in the *APC* gene lead to failure in ß-catenin degradation and consequently translocation to the nucleus and formation of constitutive ß-catenin/TCF4 complex. This finally results in aberrant intestinal gene regulation that causes a hereditary cancer syndrome called familiar adenomatous polyposis (Kinzler et al., 1991; Nishisho et al., 1991).

The classical way to explain the interaction of components of the canonical Wnt pathway in the absence and presence of Wnt ligands is the "two-state model" (Figure 0.1) (Clevers and Nusse, 2012). In the Wnt "off" state, meaning that Fz/LRP receptors are not occupied by a canonical Wnt ligand, ß-catenin is recruited to the destruction complex located in the cytoplasm, via interaction with Axin and APC. This leads kinases CK1 and GSK3 to sequentially phosphorylate ß-catenin on the more N-terminal Ser/Thr residues, which is known as the "degron" motif (Liu et al., 2002). The phosphorylated "degron" motif of ß-catenin is then recognized and bound by the E3 ubiquitin ligase F box/WD repeat protein ß-TrCP, which ubiquitinates ß-catenin and stimulates its degradation by the proteasome (Clevers and Nusse, 2012).

It was generally considered that once phosphorylated, ß-catenin left the destruction complex (Figure 0.1A) (Clevers and Nusse, 2012). However, a recent study by Clevers's group demonstrated that ß-Catenin is phosphorylated, ubiquitinated and degraded by the proteasome while still being part of the destruction complex (Figure 0.1B) (Li et al., 2012). In summary, in the Wnt "off" state, very low levels of free ß-

catenin are present in the cytosol and consequently there is no ß-catenin in the nucleus. The most well characterized nuclear interacting partners of ß-catenin are the lymphoid enhanced factor/ T-cell factor (LEF/TCF) family of transcription factors (LEF1, TCF1, TCF3 and TCF4) which provide sequence-specific binding activity (Gordon and Nusse, 2006). LEF/TCF proteins contain a high mobility group (HMG) DNA binding domain, also present in SRY/SOX transcription factors and both subfamilies recognize similar AT-rich motifs with the canonical sequence being (A/T)(A/T)CAA(A/T)GG (Clevers and van de Wetering, 1997). The consensus TCF binding sequence is AGATCAAAGG, with the core of the consensus motif being CAAAG (van de Wetering et al., 1997). In the absence of nuclear ß-catenin, LEF/TCF factors bind to Wnt target genes via their HMG domain but interact with Groucho transcriptional repressor, forming a repressive complex that prevents gene transcription (Cavallo et al., 1998; Roose et al., 1998) (Figure 0.1C).

In the "on" state (Figure 0.1), binding of Wnt ligands to the receptor complex leads to CK1-mediated phosphorylation of the cytoplasmic tail of LRP6 (Clevers and Nusse, 2012). The phosphorylated tail of LRP6 is then recognized and bound by Axin, which is the key event in Wnt signaling transduction. Higher-ordered complexes containing Dsh, Wnt and receptors have been detected (Schwarz-Romond et al., 2005) and it is considered that interaction of Dsh with the cytoplasmic tail of Fz enhances Axin-LRP interaction (Clevers and Nusse, 2012). It has been generally thought that Axin-LRP6 binding induces dissociation of the destruction complex, which leads to inhibition of ß-catenin phosphorylation and ultimately cytoplasmic accumulation of ß-catenin. However, this view was also challenged by Clevers' recent study as they found that upon Wnt signaling activation, the whole Axin destruction complex is maintained to the phosphorylated tail of LRP (Li et al., 2012). Importantly, they found that in the "on" state, ß-catenin is still phosphorylated in the destruction complex, but that Wnt inhibits ubiquitination by ß-TrCP. The authors proposed that rather to dissociate, the

complex becomes saturated with phosphorylated ß-catenin which leads to newly synthesized ß-catenin to accumulate in the cytosol (Li et al., 2012).

By a still poorly characterized mechanism, in the Wnt "on" state ß-catenin translocates to the nucleus, where it plays multiple pivotal roles in the activation of Wnt target genes (reviewed in (Gordon and Nusse, 2006)). ß-catenin forms a complex with LEF/TCF transcription factors, replacing Groucho and functioning as a transcriptional coactivator (Daniels and Weis, 2005). ß-catenin also recruits other coactivators and histone modifiers to the complex including the histone acetylase CBP/p300; CBP stands for cAMP-response-element binding protein (CREB)-binding protein (Clevers and Nusse, 2012; Gordon and Nusse, 2006). Importantly, it was shown that ß-catenin recruits the homeodomain-interacting protein kinase 2 (HIPK2) to TCF3, which phosphorylates TCF3 leading to removal of this known Wnt repressor from the promoter of Wnt target genes and to gene activation (Hikasa et al., 2010; Kim et al., 2000). It has been proposed that Tcf3 represses canonical Wnt signaling by competing with Lef1, Tcf1 and Tcf4 for nuclear ß-Catenin binding or for binding to the DNA (Solberg et al., 2012). Also interestingly, one of the direct targets of the canonical Wnt pathway is Axin2, which acts in a negative feedback loop to represses signaling by stimulating ß-catenin phosphorylation and degradation (Jho et al., 2002).

0.1.1.2 Non-canonical planar cell polarity pathway (PCP)

PCP signaling was first discovered in the fruit-fly Drosophila and has an evolutionarily conserved role in the control of cell polarization in the plane of epithelia (Seifert and Mlodzik, 2007). In vertebrates, the PCP pathway also controls mesenchymal cell migration and intercalation and is essential for shaping the gastrulating embryo (Simons and Mlodzik, 2008). Commonly referred as the non-canonical Wnt pathway, as mentionned above, PCP signaling occurs via Frizzled receptor and cytoplasmic Dsh, but does not involve ß-catenin stabilization (For

review see Simons and Mlodzik, 2008. An evolutionarily conserved core group of other transmembrane Frizzled co-receptors such as Vangl1/2, Ptk7, Celsr, Ryk, ROR2, and cytoplasmic proteins such as Prickle and Diego are essential for PCP polarizing function (Komiya and Habas, 2008). PCP signal transduction through Fz leads to activation of Dsh which then mediates activation of the small GTPases Rho and Rac (Wallingford and Habas, 2005). These in turn activate the Rho-associated kinase Rock and JNK and leads to actin polymerisation and modification of the actin cytoskeleton (Komiya and Habas, 2008). It is proposed that core members of the PCP pathway are asymmetrically distributed in the cell and that molecular polarization is transduced in cytoskeletal rearrangements essential for cell polarization and motility (Wallingford and Harland, 2001).

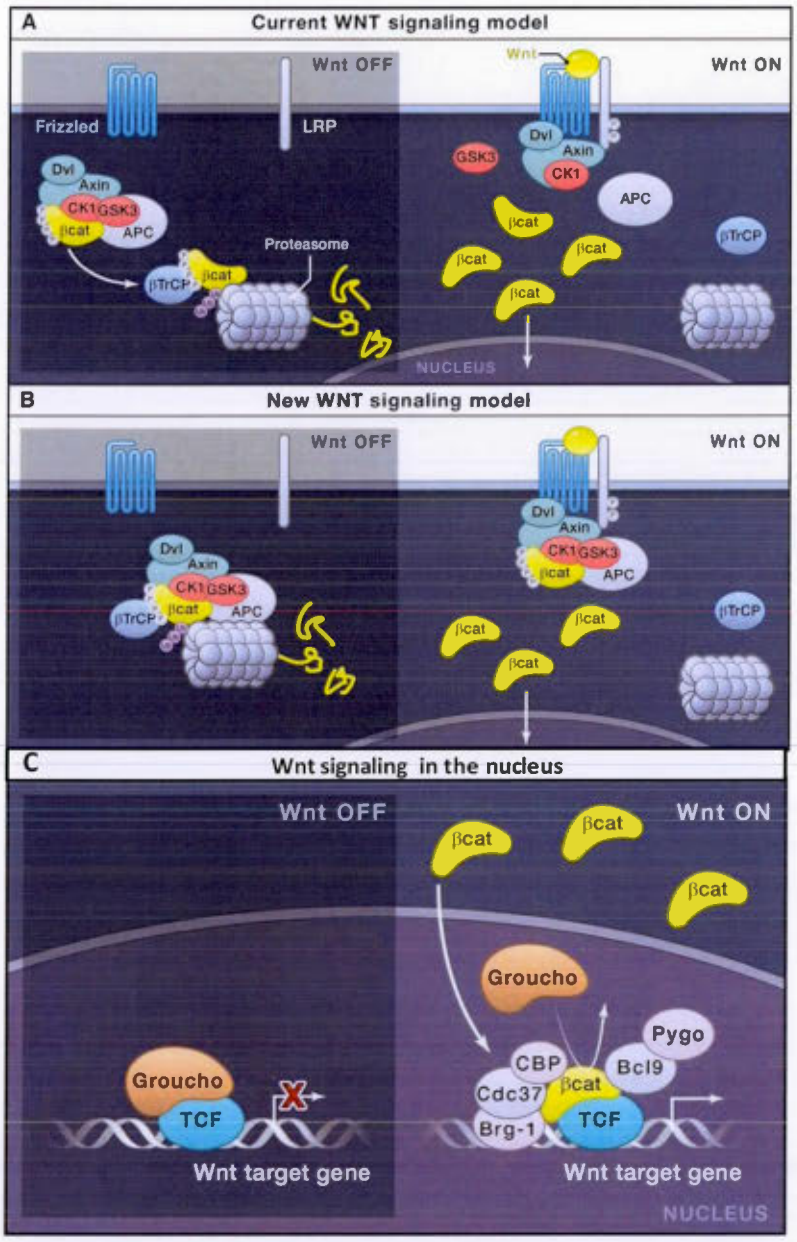


Figure 0.1 The "two state" model of canonical Wnt signaling pathway.

Taken from Clevers and Nusse, 2012. (A) Classical Wnt signaling model. In absence of a Wnt ligand (OFF state, left panel), ß-Catenin is recruited by the destruction complex and phosphorylated by CK1 and GSK3 kinases. Phosphorylated ß-Catenin leaves the destruction

complex, is ubiquitinated by the E3 ubiquitin ligase protein ß-TrCP and sent to degradation by the proteasome. In the ON state (right panel), binding of Wnt ligands to the receptor complex results in phosphorylation of the cytoplasmic tail of LRP6 by CK1,which promotes Axin-LRP interaction. This results in dissociation of the destruction complex and cytoplasmic accumulation of ß-catenin. (B) New Wnt signaling model. In the OFF state, ß-Catenin is phosphorylated, ubiquitinated and degraded by the proteasome while still being part of the destruction complex. Upon Wnt ligand binding to the Fz/LRP receptors (ON state), the destruction complex is maintained to the phosphorylated tail of LRP. ß-catenin is still phosphorylated in the destruction complex, but Wnt inhibits ubiquitination by ß-TrCP. The destruction complex then becomes saturated with phosphorylated ß-catenin and newly synthesized ß-catenin accumulates in the cytosol. (C) The two states of canonical Wnt signaling in the nucleus. In the OFF state, LEF/TCF factors bind to Wnt target genes but form a complex with Groucho transcriptional repressor that prevents gene transcription. In the ON state, ß-catenin translocates to the nucleus where it replaces Groucho and forms a complex with LEF/TCF transcription factors leading to transcriptional activation of Wnt target genes. ß-catenin also recruits other coactivators and histone modifiers to the complex.

0.1.2 Sonic hedgehog (Shh) morphogen signaling

Sonic Hedhehog (Shh) is one of the three vertebrate members of the Hedhehog (Hh) family of secreted proteins, with the other two members being Desert Hedhehog (Dhh) and Indian Hedhehog (Ihh) (Echelard et al., 1993; Krauss et al., 1993; Riddle et al., 1993). The first Hh gene was identified in Drosophila melanogaster by Nusslein-Volhard and Wieschaus (1980) in a genetic screen for mutations that affect the Drosophila larval body plan, and received its name because of the hedgehog spines-like disorganized cuticle of embryos carrying null alleles of Hh (Nusslein-Volhard and Wieschaus, 1980). Orthologs of Hh genes have also been identified in other invertebrate species including sea urchin and the cephalochordate amphioxus (reviewed in Ingham and McMahon, 2001). Molecular and functional studies of Hh in several species have evidenced their properties as intercellular signaling proteins with an evolutionarily conserved role in organization of the body plan. Hh signaling is also involved in the control of cell proliferation, survival, specification, stem cell renewal, and is – like Wnt signaling – essential for embryonic development and adult tissue homeostasis. Deregulation of Hh signaling similarly results in catastrophic

developmental defects and congenital anomalies such as holoprosencephaly as well as life threatening diseases including cancer (Briscoe and Therond, 2013).

0.1.2.1 Shh as a morphogen in the vertebrate neural tube

At the difference of Wnt that act as short range molecules (Clevers and Nusse, 2012), Shh is the classic example of a morphogen (Dessaud et al., 2008). As defined by Lewis Wolpert in its "French Flag" model, a morphogen has two main features: 1) it acts on cells at a distance from its source; and 2) acts in a concentration-dependent manner to induce differential gene expression and pattern tissues (Wolpert, 1969; Wolpert, 1996). Shh accomplishes these two requisites. Pertinent to this study and one of the best examples to illustrate the role of Shh as a morphogen in the control of tissue patterning is the vertebrate neural tube (NT), the precursor of the central nervous system (CNS) (Jessell, 2000).

During early mouse embryogenesis, Shh is sequentially produced by the node/ and its derivative the notochord, and then by floor plate (FP) cells that lie at the ventral midline of the NT (Echelard et al., 1993) (Figure 0.2A). Shh emanating from these two important organizing centers propagates to several cell diameters from the FP in the developing NT. This generate a ventral gradient of Shh activity that pattern the ventral neuroepithelium into different domains of neural progenitors: FP, p3, pMN, p2, p1 and p0 each of which generate different neuronal subtypes (Figure 0.2B) and (Dessaud et al., 2008). The identity of each progenitor domain is determined by the combinatorial expression of a group of transcriptional factors from the homeodomain and basic-loop-helix (bHLH) families that differentially respond to Shh input. Transcription factors Pax3/7 and Pax6 are repressed by Shh signaling and are conventionally known as class I genes, whereas Dbx1/2, Nkx6.1, Olig2, Nkx2.2 and Foxa2 are activated by Shh signaling and are called class II genes. Importantly, class II genes are differentially induced by a particular threshold or time of exposure to Shh signal. For example, induction of Olig2, Nkx2.2 and Foxa2, which define the most

ventral progenitors domains pMN, p3 and FP, respectively, require progressively higher concentration of Shh signal and longer time of exposure to Shh. The dorsal and ventral boundary of gene expression and thus each neural progenitor domain is then established by cross-repressive interactions between class I and class II factors, as well as between pairs of class II factors (Balaskas et al., 2012; Dessaud et al., 2008). In summary, in the ventral NT, positional information from graded and temporally changing Shh signaling is interpreted into spatial and temporal specific profiles of gene expression that subdivides the NT into different domains of neural progenitors. The combination of transcription factors expressed in a particular domain determines the identity (fate and allocation) of the neuronal subtype.

0.1.2.2 The Shh signaling pathway

Since Shh acts as a long range morphogen, the first important event to ensure Shh signaling is the effective release from the producing cell and spread to target tissues (Briscoe and Therond, 2013). Shh is synthesized as a large precursor protein of about 45 KDa that undergoes cleavage and lipid modification prior to secretion. The biologically active form of Shh is termed ShhNp and consists of an N-terminal peptide produced by autoproteolytic cleavage, which is modified with a cholesterol aduct at its C terminus and palmitoylated at its N-terminus. As they mediate multimerization as well as association in lipoprotein particles and exovesicles, both lipid modifications increase the stability of ShhNp and long range spread in the extracellular milieu. Lipid modifications are also essential for effective Shh signaling and have been proposed to mediate high affinity interaction of the protein with target receptors (Dessaud et al., 2008).

In the ventral neurorepithelium, Shh activate signaling transduction by interacting with several receptors. The first identified and perhaps best characterized is the canonical receptor Patched 1 (PTC or Ptch1), a twelve-pass transmembrane protein containing a sterol-sensing domain through which it binds Shh (Hooper and Scott,

1989; Ingham and McMahon, 2001). Other co-receptors have been shown to bind Shh and form complexes with Ptch1 that increase the binding affinity to Shh. These include the single-pass transmembrane proteins Hhip1, Cdo, Boc and Gas1 (Allen et al., 2007; Beachy et al., 2010; Dessaud et al., 2008). An important point of control of Shh signaling occurs at the level of these receptors, which not only participate in ligand binding but also are targets of Shh signaling and modulate the activity and spread of Shh. Expression of *Ptch1* and *Hhip1* is up-regulated by Shh signaling and these receptors act in a negative feedback loop or ligand-dependent antagonism (LDA) to sequester Shh ligand. Particularly Ptch1 has been shown to promote endocytosis and degradation of Shh (Incardona et al., 2002). By contrast, Cdo, Boc and Gas1 boost the Shh signaling and their expression is generally repressed by Shh signaling (Dessaud et al., 2008). Thus, receptor-mediated regulation of Shh signaling plays a key role in the spatiotemporal modulation of the Shh gradient as well as the response of cells to this gradient.

Shh signal transduction in the vertebrate NT requires the primary cilia and the activity of the transmembrane protein Smoothened (Smo), which is a member of the G protein-coupled receptor (GPCR) superfamily (Dessaud et al., 2008). Indeed, a common way to perturb Shh signaling is to use of small molecule agonists or antagonists of Smo that directly bind to the membrane-integrated heptahelical bundle of Smo (Mas and Ruiz i Altaba, 2010). One of the classical Smo inhibitors is the steroidal alkaloid cyclopamine (Chen et al., 2002). In the absence of ligand, Ptch1 represses Smo by a yet undefined mechanism (Briscoe and Therond, 2013). Ptch1 contains a sterol sensing domain and shares structural similarity to the RND bacterial transmembrane transporters. This, together with the known role of sterol-like small molecules in the modulation of Smo activity and signaling transduction, has led to the proposal that Ptch1 regulates Smo activity by transporting a putative ligand for Smo (Briscoe and Therond, 2013). Cilia also seem to play an important role in Ptch1-Smo

regulation, since in the absence of Shh, Ptch1 localizes to the cilia whereas Smo is excluded from this cellular structure (Figure 0.2C) (Dessaud et al., 2008).

Upon binding to Shh, Ptch1 releases the repression of Smo and exits the cilia leading Smo to accumulate in cilia and activate signal transduction (Figure 0.2C) (Rohatgi et al., 2007). It has also been shown that Shh stimulates phosphorylation of Smo, which induces a conformational change in the protein that facilitates its transport to the cilia and increases activity (Zhao et al., 2007). In vertebrates, Smo signaling transduction culminates in the regulation of the activity of the zinc finger transcription factors Glioma-associated oncogen (Gli) (Hui and Angers, 2011). Gli can function both as transcriptional repressors and activators, and Shh signaling results in a change in the balance between the activator and repressor forms of Gli proteins (Briscoe and Therond, 2013). There are three Gli members in amniotes: Gli1, Gli2 and Gli3, each exhibiting similar target specificity due to a high conservation in the DNA binding-zinc finger domain. The consensus Gli binding sequence is GACCACCCA (Hallikas et al., 2006). Gli2 and Gli3 are bifunctional as they contain both an N-terminal repressor domain and a C-terminal activation domain whereas Gli1 lacks the repressor domain and only acts as an activator.

How signaling is transmitted in vertebrates from Smo to the regulation of Gli is not well understood. On the basis of the knowledge of the Hh-Ci pathway (the *Drosophila* homolog of the vertebrate Shh-Gli), Smo activation involves the dissociation of a Hh-signaling complex composed by a cytosolic scaffold and kinesin-like protein named Costal 2 (Cos2), a Ser/Thr kinase Fused (Fu), the suppressor of Fu (Sufu), Gli/Ci as well as the kinases PKA, CK1α and GSK3ß (for review see Briscoe and Therond, 2013). In the absence of Shh, this complex associates with microtubules and promotes the phosphorylation of Gli/Ci by PKA, CKI and GSK3ß (Figure 0.2C). Phosphorylation of Gli2 and Gli3 targets the proteins for ubiquitination and proteolytic processing by the proteasome (Figure 0.2C). Gli2 is

almost completely degraded whereas Gli3 is converted into a repressor via removal of the C-terminal activation domain and translocates to the nucleus to represses Shh target genes (Briscoe and Therond, 2013; Dessaud et al., 2008). The activation of Smo results in dissociation of the Gli-Sufu complex and then inhibition of phosphorylation and proteolitic processing of Gli2 and Gli3, which translocate to the nucleus and activate expression of Shh target genes. Gli1 expression is activated by Shh signaling and in mammals appears to have a minor role in signaling transduction (Park et al., 2000). Mutagenesis studies in the mouse suggest that Gli2 is mainly responsible for the activation function upon Shh signaling whereas Gli3 provides most of the repressor activity (Litingtung and Chiang, 2000b; Matise et al., 1998).

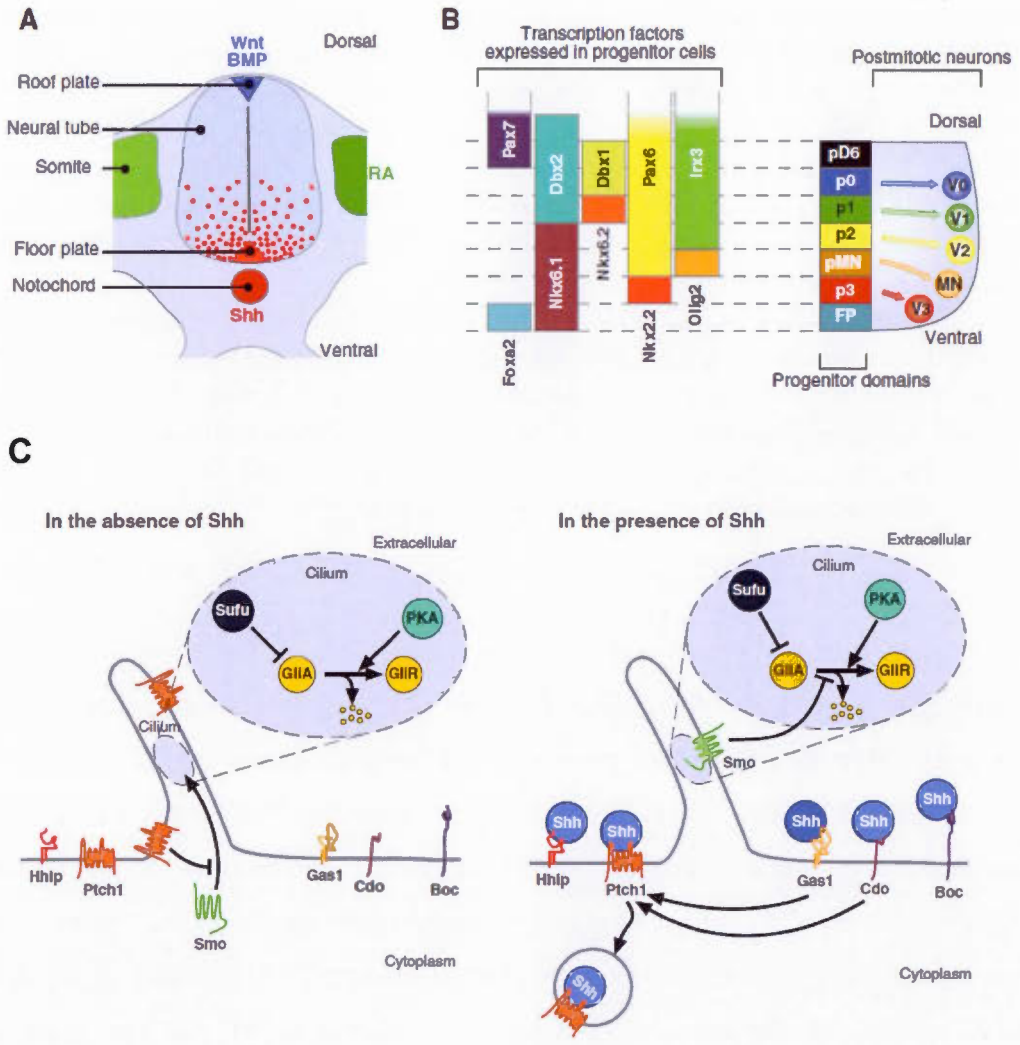


Figure 0.2 Shh morphogen signaling and the regulation of dorsal-ventral patterning in the vertebrate neural tube.

Taken from Dessaud et al., 2008. (A) Schematic of a transverse section of an amniote embryo showing the position of the neural tube, the notochord, the somites as well as the organizing centers: the floor plate (FP) and roof plate. Denoted in color are key signaling molecules involved in D-V patterning of the neural tube. Wnt and BMP molecules (blue) are produced in the dorsal neural tube. Retinoic acid molecules (RA, green) are produced by the somites. Shh molecules (in red) are secreted by the notochord and floor plate and spread along the neural tube establishing a ventral gradient of activity (red dots). (B) Schematic representation of the ventral half of the neural tube. The gradient of Shh activity controls positional identity

by regulating the expression of a particular set of transcription factors in neuronal progenitors. The differential response of these genes to graded Shh signaling and duration define their domain of expression and then the identity of the neural progenitor domain: FP, p3, pMN, p2, p1 and p0. Each progenitor domain generates different ventral (V) interneuron subtypes (V0-V3) or motoneurons (MN). (C) Shh signaling pathway in the vertebrate primary cilia. In absence of Shh (left panel), the receptor Patched 1 (Ptch1) represses Smoothened (Smo) translocation to the cilia. Gli2 and Gli3 transcription factors are phosphorylated by the protein kinase A (PKA), targeted to ubiquitination and proteolytic processing by the proteasome. Gli2 is almost completely degraded whereas Gli3 is converted into a repressor and translocates to the nucleus to represses Shh target genes. Binding of Shh to Ptch1 (right panel) releases the repression of Smo which accumulates in cilia. Other co-receptors including the single-pass transmembrane proteins Hhip, Cdo, Boc and Gas1 also bind Shh and form complexes with Ptch1 that increase the binding affinity to Shh. In cilia, Smo inhibits phosphorylation and proteolitic processing of Gli2 and Gli3, which translocate to the nucleus and activate expression of Shh target genes.

0.1.3 Canonical Smad-dependent BMP signaling

Bone Morphogenetic Proteins (BMPs) are multifunctional secreted glycoproteins belonging to the TGF-ß superfamily (Carreira et al., 2014). As their name suggests, BMPs were first identified as potent inducers of bone formation (Reddi and Huggins, 1972; Urist and Strates, 1971), and constitute interesting therapeutic candidates in bone and cartilage tissue repair (Carreira et al., 2014). Besides these important roles, BMPs play key pleiotropic roles in embryonic development, adult tissue homoeostasis and disease (Hogan, 1996; Miyazono et al., 2010; Walsh et al., 2010). BMPs are major orchestrators of embryonic patterning of the mesoderm and neuroectoderm (De Robertis and Kuroda, 2004). They notably regulate neural, limb, tooth, kidney, skin, hair, muscle and hematopoietic organ development (Herpin and Cunningham, 2007; Miyazono et al., 2010). In adulthood, BMPs control iron metabolism and homeostasis of the vascular system (Miyazono et al., 2010).

The BMP family is composed of about 20 members in human, classified in four subfamilies based on amino acid sequence similarity (reviewed in Carreira et al., 2014). For example, BMP2 and BMP4 share 80% of homology and are in the same

subgroup (A), whereas BMP 5, 6,7,8 are in the subgroup C, also known as OP-1 (from osteogenic protein-1 subgroup) (Carreira et al., 2014; Miyazono et al., 2010). BMPs are small glycoproteins of 120 amino acids in size that contain several conserved cysteine residues that are essentials for their activity, as they mediate homo and hetero-dimerization of the protein via disulfide-links (Walsh et al., 2010). Like many growth factors, BMPs are synthesized and secreted as a large inactive precursor. Particularly, they carry an N-terminal latency/signal hydrophobic peptide that is proteolitycally cleaved in the extracellular space by specific proteases (Constam and Robertson, 1999) (Figure 0.3). The released mature protein dimerizes to form the bioactive form that binds to a cell surface complex of type I (BMPR1) and type II (BMPR2) serine/threonine kinase receptors in the target tissues (Figure 0.3) (Carreira et al., 2014; Miyazono et al., 2010). The BMPR1 group is subdivided into type 1a (BMPR1a also known as ALK3) and type 1b (BMPR1b also known as ALK6) (Miyazono et al., 2010). Type II receptor kinase is constitutively active and phosphorylated (Carreira et al., 2014).

Upon BMP binding and oligomerization, type II receptor recruits and transphosphorylates the Type I receptor, activating its kinase function (Figure 0.3) (Carreira et al., 2014; Herpin and Cunningham, 2007). Activated type I receptor subsequently propagate the signal via recruitment and phosphorylation of receptor-activated (R) Smad proteins (R-Smad) 1/5/8, in a canonical pathway, or via several tyrosine kinases in a Smad-independent non-canonical pathway (Herpin and Cunningham, 2007; Moustakas and Heldin, 2009). Other co-receptors and scaffold proteins act to facilitate interaction of R-Smad proteins with the receptor complex (Walsh et al., 2010). In the canonical pathway, phosphorylated-Smad1/5/8 (pSmad1/5/8) binds to a co-Smad protein (Smad4) in the cytosol, forming a complex that translocates to the nucleus and regulates target gene expression. The consensus binding sequence for Smad transcription factors in target promoters has been defined as 5'- GTCT-3' (Jonk et al., 1998). Others have reported that this is not a high

affinity binding element and have rather shown Smad binding to *Drosophila* Mad-like consensus sequence 5'-GCCGnCG-3'(Alvarez Martinez et al., 2002).

The differential response triggered by different BMPs molecules and other members of the TGF-ß subfamily (including nodal, activin, inhibin, and other factors) is mediated by the interaction with specific type I and II receptor and activation of different R-Smad proteins (Herpin and Cunningham, 2007; Miyazono et al., 2010; Walsh et al., 2010). For example, it has been shown that BMP 2, 4 and 7 preferentially signal via Smad 1, 5 and 8, whereas other TGF-ß members such as nodal particularly induce activation of Smad 2 and 3. This is modulated by the specificity and affinity of binding of TGF-ß members to a particular type I receptor. For example, it has been shown that members of the BMPR1 group (ALK-3 and ALK-6), which are strongly bound by BMP2 and BMP4, preferentially induce activation of Smad 1, 5 and 8, whereas the TßR-I group (ALK4, 5, 7) activate Smad2/3 (reviewed in (Miyazono et al., 2010)).

The output of BMP signaling is modulated at several levels (Figure 0.3). Synthesis of BMP molecules is controlled at the epigenetic level, particularly by methylation of BMP gene promoters (Wen et al., 2006). The strength and extent of BMP signaling transduction is controlled by phosphatase-dependent dephosphorylation of type I receptor and R-Smad and also by inhibitory Smads, particularly Smad 6 and 7, which compete with Smad1/5/8 for receptor binding (Imamura et al., 1997; Nakao et al., 1997). One of the most precise and well studied points of control of BMP signaling occurs at the extracellular space, by the action of BMP antagonists (Carreira et al., 2014; Walsh et al., 2010). These are secreted proteins that are also produced in an inactive form containing a latency domain (Avsian-Kretchmer and Hsueh, 2004). Proteolytic cleavage of the precursor forms a mature protein that directly binds and sequesters BMP molecules (Groppe et al., 2002; Walsh et al., 2010).

To date, over 15 BMPs antagonists have been identified and are classified in four groups: (a) Dan or Neuroblastoma 1, including Gremlin, Cerberus and others; (b) Twisted gastrulation (Tsg); (c) Chordin family (including Chordin, Noggin, Follistatin and others) and (d) Follistatin-related gene (FLRG) (Carreira et al., 2014; Walsh et al., 2010). Because of their potency, expression of BMP antagonists is tightly controlled during development, with aberrant expression of BMP inhibitors causing several drastic developmental defects such as holoprosencephaly and head truncation in mouse, and diseases such as fibrosis and cancer (Bachiller et al., 2000; Walsh et al., 2010).

Studies in *Xenopus* have shed light into the mechanisms of action of the BMP inhibitors. These studies show that Chordin and Noggin directly bind with high affinity to BMP2 and BMP4 in the extracellular space and block the interaction of BMPs to type I and II receptors (Piccolo et al., 1996; Zimmerman et al., 1996). Noggin also antagonizes BMP7 (Groppe et al., 2002; Zimmerman et al., 1996). Another signaling molecule, Follistatin, exhibits wider specificity for members of the TGF-ß superfamily and inhibits Activin and also BMP7 and BMP4 (Liem et al., 1997; Yamashita et al., 1995). In the mouse, *Chordin* and *Noggin* are co-expressed in the anterior primitive streak (PS), node and axial mesendoderm derivatives (Bachiller et al., 2000; McMahon et al., 1998). Comparison from single and compound mutants phenotypes suggest that these anti-BMP factors function redundantly in the control of dorsal-ventral (D-V) patterning (Bachiller et al., 2000; McMahon et al., 1998).

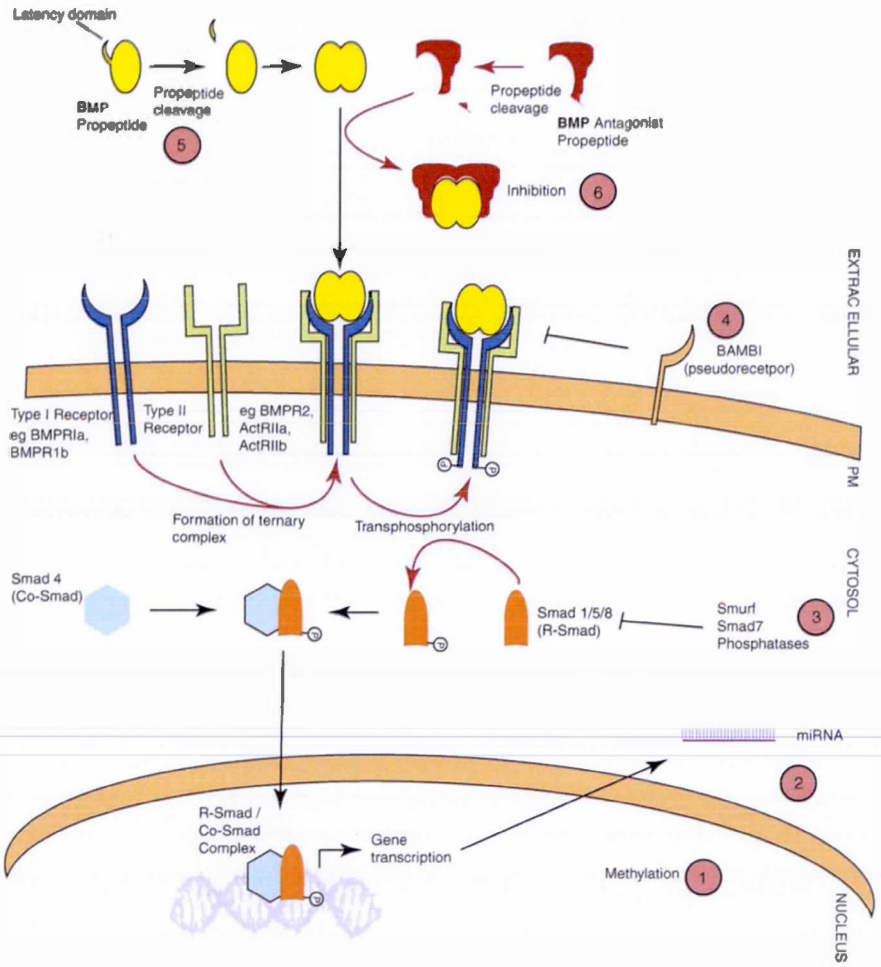


Figure 0.3 Canonical BMP-Smad signaling pathway.

Taken from Walsh et al., 2010. BMPs are synthesized and secreted as a large inactive precursor bearing an N-terminal latency domain that is cleaved in the extracellular domain by specific proteases. Proteolitycally cleaved BMP monomers form active dimers and bind to a cell surface complex of type I (BMPR1) and type II (BMPR2) serine/threonine kinase receptors. Upon receptor binding and oligomerization, type II receptor recruits and transphosphorylates the Type I receptor which subsequently propagate signal via recruitment and phosphorylation of receptor-activated (R) Smad proteins (R-Smad) 1/5/8. Phosphorylated R-Smads form a complex with the co-Smad (Smad4) in the cytosol and translocate to the nucleus to regulate target gene transcription. Steps 1-6 represent different levels of intracellular and extracellular modulation of BMP signal transduction.

0.2 Mouse embryogenesis

As described by Hogan, the exquisitely beautiful form of a developing embryo is sculpted from one single cell – the fertilized egg – and requires the coordination of morphogenetic movements, cell division, proliferation, migration, differentiation and genetically programmed cell death (Hogan, 1999). The architecture of the developing embryo as well as the rate of embryogenesis considerably differs between vertebrate species. Whereas frog embryos complete development within a few days, mammalian embryogenesis is very slow, taking 19-20 days (3 weeks) in the mouse and around 40 weeks in human. The objective of this chapter is to provide basic elements that help to understand how the mouse embryo develops in the first third of the gestation period, from embryonic day (e) 0.5 to 7.5. The first week of embryogenesis is perhaps the most important time in a mouse's life. In this short period the three definitive embryonic germ layers – ectoderm, mesoderm and endoderm – are allocated and the basic body plan sculpted with the specification of anterior-posterior (A-P) and D-V body axes. After this time, tissue diversity and allocation is established by regionalization or patterning into increasingly smaller developmental fields.

0.2.1 The first week of mouse embryogenesis

As represented in Figure 0.4, in the first week of embryogenesis the mouse embryo progresses through several stages from fertilization to gastrulation (Arnold and Robertson, 2009; Beddington and Robertson, 1999; Nagy et al., 2003; Rossant and Tam, 2009; Takaoka and Hamada, 2012). During the first three days of development (e0.5 –e3), the small fertilized egg undertakes serial rounds of cleavage divisions to form two, four, eight and sixteen cells, which are called blastomeres. In this short period, known as the cleavage stage, the maternal mRNA is degraded and the zygotic genome is activated; also the eight-cell embryo begin to experience compaction and polarization (Rossant and Tam, 2009). At e3 the sixteen-cell embryo looks like a

mulberry. Hence, this stage is frequently called the morula stage (from the latin *morus*, meaning mulberry). Asymmetrical cell divisions lead to polarization and segregation of cells into the first two subpopulations of the embryo: the inner cell mass (ICM) and the trophectoderm (TE). Cells located on the inside of the morula are apolar and give rise to the ICM, whereas cells located outside are polarized and give rise to the TE (Johnson and Ziomek, 1981). This lineage segregation is controlled by a reciprocal feedback mechanism between the transcription factors Cdx2 and Oct3/4, which are exclusively expressed in the TE and ICM, respectively. Cdx2 is required for TE expansion, whereas Oct3/4 is required for maintenance of the pluripotency of the ICM (Niwa et al., 2005).

In the subsequent step, recognized as the blastula stage (e3.5), the morula embryo undertakes cavitation, becoming a hollow vesicle known as the blastocyst (Beddington and Robertson, 1999; Rossant and Tam, 2009). The blastocyst is composed of an external layer of TE cells, a fluid-filled cavity named the blastocoel, and the ICM (Figure 4A). The blastocyst is polarized, with the ICM situated at its proximal pole. At around e4.0, the layer of ICM cells that lies close to the blastocoel differentiates into the primitive endoderm (PrE) or hypoblast, whereas the remaining ICM cells become the epiblast (also called primitive ectoderm). The PrE and the epiblast exhibit mutual exclusive expression of the GATA-binding transcription factors GATA6 and Nanog, respectively (Lanner and Rossant, 2010; Yamanaka et al., 2010).

At late blastocyst stage (e4.5) the three primary tissue types are formed: the TE, epiblast and PrE (Arnold and Robertson, 2009; Beddington and Robertson, 1999; Gardner, 1978; Gardner, 1983). These are progenitors of all tissues of the embryo, as well as extraembryonic tissues that support embryo development. The TE gives rise to trophoblast cells and extraembryonic ectoderm (ExE) that forms the ectoplacental cone and the chorion component of the placenta. The PrE will form the

extraembryonic parietal endoderm, that directly contact the maternal tissue, as well as the visceral endoderm (VE) which gives rise to the endoderm of the visceral yolk sac. The epiblast contributes to all germ layers of the embryo proper including the germ line as well as to the extraembryonic mesoderm (Gardner, 1978). The later contributes to the amniotic mesoderm, allantois and chorionic mesoderm, as well as hematopoietic precursors. Pluripotent mouse embryonic stem (ES) cells can be derived from the ICM/epiblast of pre-implantation morula/blastocyst embryos and used to form chimeras (Brook and Gardner, 1997).

The cleavage, morula and early blastula stages take place while the embryo is moving along the oviduct to the uterus (Nagy et al., 2003). At e4.5, the blastocyst implants in the uterus. The TE external layer mediates this process, with TE cells that surround the blastocoel being the first to adhere to the antimesometrial uterine wall. After implantation, at around e5, the TE closer to the ICM expands and forms the extraembryonic ectoderm (ExE). At the proximal pole of the embryo the ExE forms the ectoplacental cone and a cup-shaped layer of epithelial cells. It has been proposed that expansion of these tissues pushes the epiblast and the enveloping visceral endoderm to the distal pole (Beddington and Robertson, 1999). This mechanical pressure, together with a combination of apoptotic and survival signals from the visceral endoderm forms a cavity in the center of the epiblast known as the proaminiotic cavity (Coucouvanis and Martin, 1995). The epiblast then becomes a cup-shaped epithelium and the embryo takes the shape of an "egg cylinder". From e6.0 onwards, epiblast cells rapidly and extensively proliferate to support the changes in morphology and cell arrangement that take place during the next developmental process: gastrulation.

The combination of fate specification and morphogenetic movements transforms the epiblast into three embryonic germ layers: the ectoderm, mesoderm and definitive endoderm (Arkell and Tam, 2012; Arnold and Robertson, 2009). This process is

known as gastrulation, and is a key event in embryogenesis that establishes the basic body plan and organ primordia of the future mouse. As the British embryologist Lewis Wolpert described in his famous phrase (1986): "It is not birth, marriage, or death, but gastrulation, which is truly the most important time in your life".

Gastrulation begins at e6.5 with formation of a structure called the primitive streak (PS) by the convergence of epiblast cells in the proximal posterior part of the epiblast, close to the embryonic-ExE junction (Lawson et al., 1991) (Figure 0.4A). The location of the PS marks the posterior part of the embryo, and the opposite region is considered the anterior side by default (Arnold and Robertson, 2009; Beddington and Robertson, 1999). Epiblast cells that ingress through the PS experience an epithelial to mesenchymal transition (EMT) and form the nascent mesoderm which lies between the epiblast and the visceral endoderm (Figure 0.4B). BMP, Nodal and canonical Wnt signaling are essential for mesoderm formation (reviewed in (Arnold and Robertson, 2009)). As gastrulation proceeds, the PS elongates and extends to the distal tip of the embryo.

The epiblast region that surrounds the PS exhibits extensive cell proliferation. New progenitors are constantly generated from this zone. The progressive addition of new tissue at the posterior end of the embryo results in elongation of the body axis. Gastrulation is thus a constant process in the posterior region that last around 6 days (from e7.5 to e13.5 approximately) (Young et al., 2009). Axial elongation depends on the maintenance of a posterior stem cell zone, which constantly generate a population of axial stem cell-like progenitors. Clonal analysis has traced this population to the node-streak border and anterolateral PS region (Cambray and Wilson, 2007). Fate mapping analyses suggest that axial stem cell-like progenitors are bipotent and can give rise to neural and mesoderm progenitors (Cambray and Wilson, 2007). Concomitantly with axial elongation, the node and PS regress caudally until disappearing. The founder function of the PS is replaced by its descendant the tail

bud. Transplantation studies have suggested that a particular region of the tailbud known as the chordoneural hinge (CNH) is the source of axial stem cell progenitors (Cambray and Wilson, 2002).

Since new tissue is constantly added from the posterior region (PS/tailbud) during axial elongation, it is important to note that mammalian embryos develop in an anterior to posterior order. This means that the most anterior part of the embryo – the head – is the oldest-one whereas the most posterior one – the trunk and tail – is the youngest. Consequently, tissues located more anteriorly differentiate before the most posterior-ones. This is well exemplified by the process of somitogenesis, during which blocks of paraxial mesoderm known as somites are generated in a rostral to caudal sequence (Pourquie, 2003).

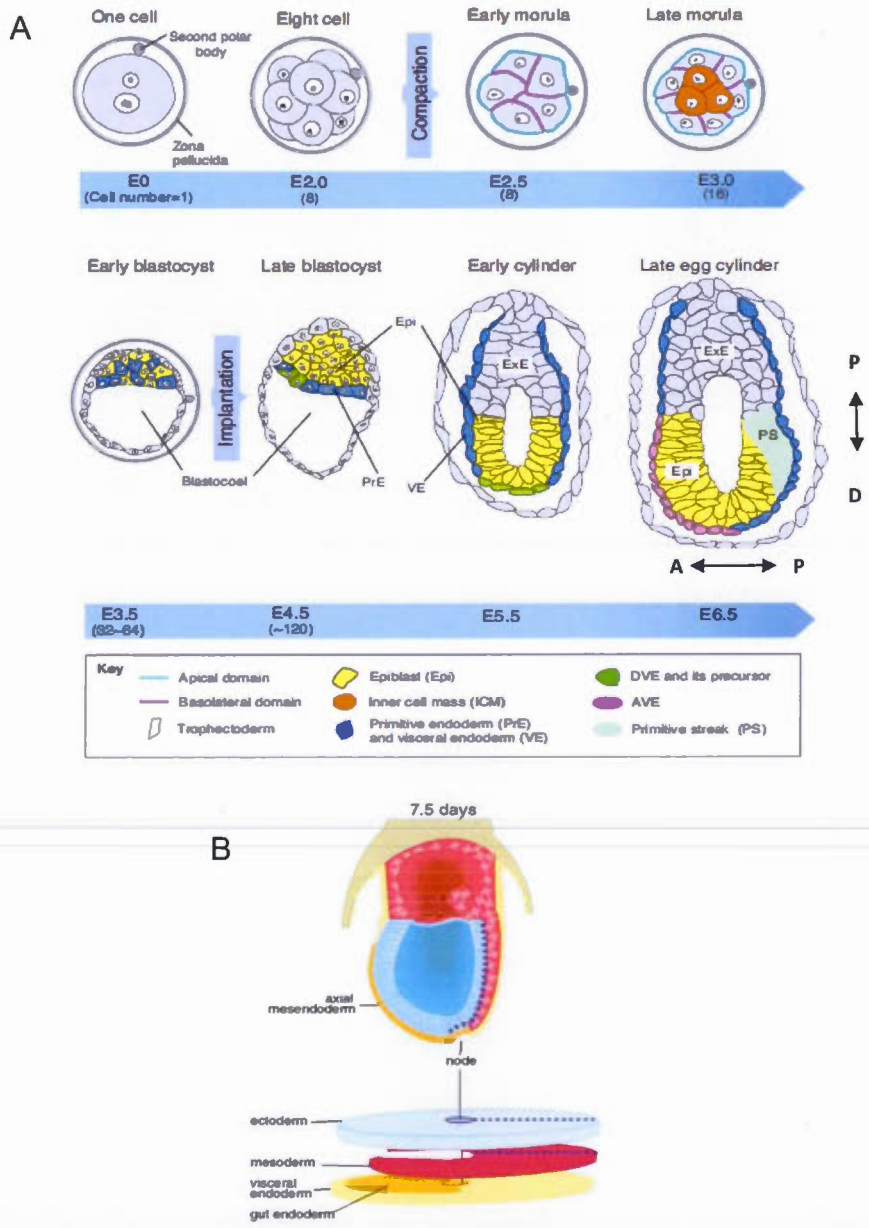


Figure 0.4 The first week of mouse embryogenesis.

(A) Representation of the major events that take place from fertilization to gastrulation. Taken from Takaoka and Hamada, 2012. On the top is the preimplantation development of the embryo from fertilization to embryonic day (e) 3.0. On the bottom are the subsequent steps from early blastocyst stage (e3.5) to gastrulation (e6.5). AVE: anterior visceral endoderm; DVE: distal visceral endoderm; Epi: epiblast; ExE: extraembryonic ectoderm; PrE: primitive endoderm; PS: primitive streak; VE: visceral endoderm. P: proximal; D: distal. A: anterior; P: posterior. (B) Representation of a gastrulating e7.5 mouse embryo showing

the relative position of the node, PS, and the embryonic germ layers. Taken from Beddington and Robertson, 1999.

0.2.2 Establishment of the Anterior-Posterior and Dorsal-Ventral axes

The mouse body, as well as that of all vertebrates, is asymmetric. The anterior (head) is morphologically different from the posterior (trunk and tail); the back (dorsal) is different from the belly (ventral). Although less evident, the left part of the body is also different from the right, with some visceral organs for example the heart and liver located at one side of the body. Body asymmetry is established early during embryonic development by the work of organizer centers that polarize signaling cues and gene expression patterns. This leads to formation of first the anterior-posterior (A-P), then the dorsal-ventral (D-V) and finally the Left-Right (L-R) axes. The three major axes coordinate cell identity (fate and allocation) during embryogenesis, and thus build the complex structure of the body. This section is dedicated to the brilliant British embryologist Rosa Susan Penelope Beddington (1956-2001) who made a major contribution to the understanding of the mechanisms that control axial patterning in the early mouse embryo. For excellent reviews on this topic, see (Arnold and Robertson, 2009; Beddington and Robertson, 1999; Takaoka and Hamada, 2012; Tam and Behringer, 1997).

0.2.2.1 Establishment of the A-P axis

Formation of the PS at gastrula stage (e6.5) is the first morphological sign of A-P asymmetry. However, studies from Beddington's lab demonstrated that the e5.5 pre-gastrula embryo, while apparently symmetric, is already patterned and exhibits differential gene expression along its proximal and distal poles (Thomas et al., 1998). Indeed, this study showed that the homeobox gene *Hex* is uniquely expressed in a few visceral endoderm cells located at the distal tip of the e5.5 mouse embryo. *Hex*-positive distal VE cells are now recognized as the DVE center (from distal visceral endoderm) (Figure 0.4A). Moreover, by using DiI labelling techniques the authors

showed that the DVE moves unilaterally, at e6, towards the prospective anterior domain of the embryo and forms the anterior visceral endoderm (AVE). This led to propose a model in which the anterior-posterior axis is the consequence of the rotation of the pre-established proximal-distal axis (Beddington and Robertson, 1999; Thomas et al., 1998).

More recently, Hamada's group traced back the origin of the DVE cells to the late blastocyst stage and re-challenged the idea that AVE is derived from the DVE (Takaoka et al., 2011). Indeed, this study showed that the DVE marker left-right determination factor 1 (*Lefty1)* is already expressed in a subset of epiblast cells and then in the PrE progenitors as early as e4.2. Moreover, the authors showed that AVE is not entirely derived from *Lefty 1* positive DVE cells, as the removal of these cells did not affect AVE formation. Instead, the authors proposed that relocation of the DVE generates global movements of VE cells and newly induced AVE from *Lefty1* negative VE cells that migrate to the distal tip of the embryo at e5.5. Then, AVE cells and DVE cells are two separated populations; AVE follows DVE migration and finally occupies the entire prospective anterior pole (Figure 0.4A).

Work from Beddington's and Robertson's groups (as well as other labs) has yielded insights into the molecular mechanisms that initiate proximal-distal polarity (reviewed in Arnold and Robertson, 2009), and has demonstrated crucial roles of Nodal (a member of the transforming growth factor–ß (TGF-ß) family) in this process (Brennan et al., 2001). It has been shown that Nodal induces formation of the DVE by activating a Smad2-dependent regulatory cascade that leads to the expression of genes coding for transcription factors such as Hex, FoxA2, LIM1, Otx2, Gsc, and secreted molecules such as Lefty 1, Cerberus-like protein 1 (Cer1), and Dickkopf homologue 1 (Dkk1) (Perea-Gomez et al., 2001; Perea-Gomez et al., 1999; Thomas et al., 1998). Consequently, the DVE secretes the Nodal antagonist Lefty1 and Cer1

as well as the canonical Wnt antagonist Dkk1, and therefore inhibits Nodal and Wnt inputs in the overlying distal epiblast.

In the opposite side of the epiblast, close to the ExE, nodal acts by a Smad2-independent mechanism to reinforce its own input and crosstalk with BMP4 signaling from the ExE as well as canonical Wnt in the proximal epiblast to also intensify these inputs. Hence, Nodal patterns the pre-gastrula embryo by establishing differential gene expression and signaling along the proximal-distal axis. The proximal epiblast exhibit high levels of posteriorizing Nodal and canonical Wnt signaling and is fated to become the posterior pole. The distal epiblast does not receive these signals due to the repressive action of Lefty1, Cer1, Dkk1 secreted from the DVE center, and is fated to become the anterior pole. Nodal also controls migration of the DVE to the prospective anterior pole, with the Nodal-Smad2 target Otx2 and the proximal-distal gradient of Nodal signaling being essential in this process (Perea-Gomez et al., 2001; Yamamoto et al., 2004).

The AVE, also known as the head organizer, secretes antagonists of Nodal, BMP and Wnt signaling, and prevents formation of the PS in the nearby epiblast (Perea-Gomez et al., 2002). Thus, location of the AVE establishes the anterior pole of the embryo, the head, and specifies the prospective anterior epiblast to neuroectoderm. On the opposite side of AVE at e6.5, integration of high levels of nodal and Wnt-ßcatenin signaling in the proximal posterior epiblast with BMP signaling from the ExE induces the formation of the PS and sets the posterior pole of the embryo (Ben-Haim et al., 2006; Fuentealba et al., 2007; Zeng et al., 1997). These signals induce nearby epiblast cells to ingress through the PS and thus to be fated to mesodermal or endodermal lineages, as described before.

The A and P poles are separated by the node, which as mentionned before is located at the most anterior end of the PS. At e7.5, the embryo is evidently patterned with the anterior pole expressing the anterior markers Otx2 (neuroectoderm), FoxA2

(mesendoderm) and also the Nodal and Wnt inhibitors Cer1 and Dkk1. By contrast, the posterior region is characterized by the expression of posterior markers such as Brachyury (*T*) expressed in the PS, node and notochord, Cdx (ectoderm and mesoderm) and Wnt3a. Expression patterns can be found in (Arnold and Robertson, 2009). Products of these genes will subsequently control axial identity and initiate cell fate commitment.

0.2.2.2 Establishment of the D-V axis

With the formation of the proamniotic cavity at the "egg cylinder" stage (e5), the prospective dorsal and ventral poles of the embryo become evident. The region of the epiblast underlying the proamniotic cavity marks the prospective dorsal pole whereas the outer surface of the VE denotes the prospective ventral pole (Tam and Behringer, 1997). However, the proper D-V patterning in terms of tissue allocation, signaling polarization and gene expression pattern is established at gastrulation after the A-P axis is formed and depends on a signaling center known as the organizer (Beddington and Robertson, 1999; De Robertis and Sasai, 1996).

The major molecular event responsible for D-V axis formation in vertebrates is the polarization of BMP signaling by the action of BMP antagonistic extracellular molecules such as chordin, noggin, follistatin (Bachiller et al., 2000; De Robertis and Kuroda, 2004; Harland and Gerhart, 1997; McMahon et al., 1998; Sasai et al., 1994; Smith and Harland, 1992). The anti-BMP/ TGF-ß regulatory network is conserved in non-vertebrate Bilaterian species like Drosophila, with the BMP homologue being Decapentaplegic growth factor (Dpp), and the anti-BMP chordin homologous being Short gastrulation (sog) (De Robertis and Sasai, 1996). Mutations in members of these pathways in Drosophila, *Xenopus*, zebrafish and mouse affect patterning of the neuroectoderm and dorsal mesoderm, clearly defining the key role of BMP polarization in the establishment of the D-V axis (Bachiller et al., 2000; De Robertis and Sasai, 1996; Hammerschmidt et al., 1996). For example, mice double

homozygous mutants for *Chordin* and *Noggin* exhibit an increase in BMP signaling that result in holoprosencephaly as well as notochord and somite defects, among other phenotypes (Bachiller et al., 2000). This means that formation of dorsal tissues like the neuroectoderm, somites and notochord requires inhibition of BMP signaling.

The organizer

Whereas specification of the A-P axis depends on the AVE, establishment of the D-V axis is controlled by the organizer. This cell population present during gastrulation is the source of many BMP antagonists – as well as Wnt inhibitors – and is thus responsible for the spatiotemporal inhibition of BMP signaling (Harland and Gerhart, 1997; McMahon et al., 1998). The organizer is required for neural induction and D-V patterning. The name organizer comes from pioneer studies performed by Hans Spemann and Hilde Mangold in 1924 using salamander gastrula eggs. These embryologists grafted a region located in the dorsal mesoderm of the gastrula embryo (known as the dorsal blastopore lip) to a region at the ventral side of another embryo that normally would form epidermis. Surprisingly, the host embryo had two complete body axis including two heads (Hemmati-Brivanlou and Melton, 1997). This experiment clearly showed that the donor tissue – many years later named Spemann's organizer (Harland and Gerhart, 1997) – had all the information necessary to induce a dorsal axis including a complete secondary neuraxis and other dorsal axial structures (Spemann, 1924). Similar organizers were then found in other vertebrate species including chick, rabbit, and mouse. The chick organizer is named Hensen's node (Waddington, 1933; Waddington, 1936). The mouse organizer is called the gastrula organizer (GO) and forms at the anterior end of the PS during gastrulation (Beddington, 1994; Beddington and Robertson, 1999).

Combined cell grating, dye labeling and live imaging analysis have revealed that the mouse GO is a dynamic cell population that progress from the Early- and Midgastrula-Organizer (EGO and MGO) to the node (trunk organizer) (Kinder et al.,

2001). EGO and MGO exhibit earlier organizing properties, and give rise to anterior mesendoderm (AME) derivatives such as the head process notochord that organize and pattern the anterior neuraxis (Yamanaka et al., 2007) (Figure 0.5A). The node acts at the late PS stage in the organization of posterior neural development and gives rise to the posterior notochord (Yamanaka et al., 2007). The GO-derived axial mesendoderm derivatives such as the notochord extend along the A-P axis and secrete BMP antagonists that restrict BMP activity to the dorsal pole. The notochord also secretes the ventralizing morphogen Shh. Counteracting gradients of dorsalizing BMP and ventralizing Shh signals create a D-V axis in the NT and somites (Bachiller et al., 2000; Marcelle et al., 1997; McMahon et al., 1998; Tanabe and Jessell, 1996).

D-V patterning in the mouse

In the gastrulating mouse embryo, BMP signaling is restricted to the proximal posterior region (Arnold and Robertson, 2009; Yang and Klingensmith, 2006) (Figure 0.5B). BMP signaling exhibits ventralizing activity, so epiblast derivatives that receive BMP signaling will locate at the ventral pole, whereas those that do not receive this signal will lie dorsally. Whether or not cells receive BMP activity is determined by their proximity to the GO, the source of BMP inhibitors. The future D-V location of epiblast derivatives is also determined by whether cells ingress through the PS or not, as well as by the site of ingression (Tam and Behringer, 1997) (Figure 0.5C). Cells that do not ingress in the PS are specified into ectoderm which subsequently divides into dorsal and ventral ectoderm. The dorsal ectoderm is protected from the ventralizing action of BMP by the coordinated action of AVE, EGO and MGO, and gives rise to the neuroectoderm which subsequently forms the nervous system. The ventral ectoderm receives BMP signaling and gives rise to epidermis. The mesendoderm is also patterned in dorsal and ventral mesoderm according to their accessibility to BMP signaling (De Robertis and Kuroda, 2004). Progenitors that enter through more posterior regions of the PS and thus are exposed to BMP signaling develop into ventral mesoderm derivatives such as the lateral plate

mesoderm, whereas those that enter through more anterior regions of the PS, are closer to the GO – and thus receive BMP antagonism – and give rise to dorsal mesoderm derivatives that will further develop into somites and notochord (Figure 0.5B). This Dorsal (i.e: CNS, dorsal mesoderm)- Ventral (i.e.: lateral plate mesoderm, ventral blood islands) plan is conserved in all vertebrates, and the molecular mechanisms that establish it are conserved as well (De Robertis and Kuroda, 2004; De Robertis and Sasai, 1996; Tam and Behringer, 1997).

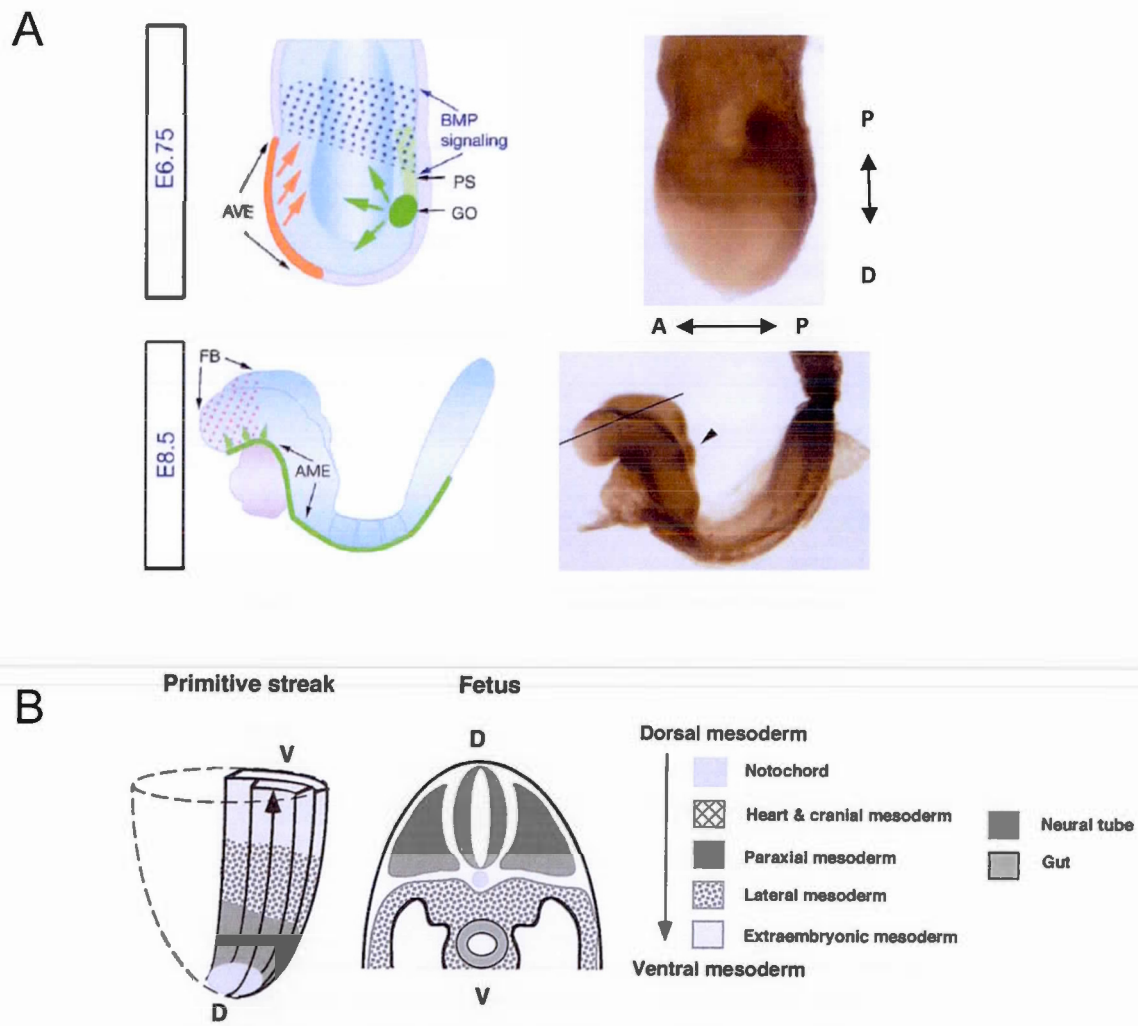


Figure 0.5 BMP signaling and Dosal-Ventral (D-V) patterning of the early mouse embryo

(A) Distribution of organizing centers and BMP signaling in the mouse embryo from e6.75 - e8.5 according to immunostaining for pSmad1/5/8. Adapted from Yang and Klingensmith, 2006. At e6.75, the gastrula organizer (GO), and the anterior visceral endoderm (AVE) restrict BMP signaling to the proximal epiblast. The anterior mesendoderm (AME) derivatives of the GO organize and pattern the anterior neuraxis at e8.5 by maintaining BMP antagonism. Note that at this stage BMP signaling activity is restricted to the lateral borders of the anterior and posterior neural folds. P: proximal; D: distal. A: anterior; P: posterior. (B)

Cell lineage allocation according to the D-V axis. Note the relationship between the site of ingression of epiblast cells through the PS and their future D-V location in the fetus. Adapted from Tam and Behringer, 1997.

0.3 Transducing positional information into cell identity

Formation of the A-P and D-V axis clearly polarizes signaling cues and establishes the basic body plan. The next obvious question is how signaling information is transduced at the molecular level to pattern initial homogeneous populations of mesodermal, endodermal and ectodermal cells into the organized diversity of tissues and structures in the organism? It has been proposed that cell fate or patterning is determined by the integration of signaling cues at a relative axial position (Deschamps and van Nes, 2005).

0.3.1 The Hox code and A-P patterning

Considerable work has now established that posteriorizing signals from fibroblast growth factor (FGF), retinoic acid (RA) and canonical Wnt pathways impart A-P positional information via regulation of *Hox* genes and particularly via establishment of a molecular "*Hox* code" consisting of distinct combinations of *Hox* gene products in specific axial levels at a given time (Deschamps and van Nes, 2005; Kessel and Gruss, 1991).

Hox genes encode homeodomain transcription factors, relative of the *HOM-C* genes of *Drosophila* (Brooke et al., 1998). In the mouse genome there are 39 *Hox* genes distributed in four clusters: *Hoxa-Hoxd* (Figure 0.6A). Based on their physical position within each cluster as well as similar spatiotemporal expression patterns and function, *Hox* genes are arranged in 13 paralogous groups (although not all clusters have 13 paralogs) (Krumlauf, 1994). Murine *Hox* expression starts at e7.25 in the posterior PS and extends anteriorly in the neuroectoderm and mesoderm to reach a characteristic anterior boundary (Figure 0.6B), which is located more anteriorly in

neural tissues relative to mesoderm (Figure 0.6C) (Deschamps et al., 1999; Deschamps and van Nes, 2005).

A special hallmark of the *Hox* clusters is the correlation between physical chromosomal organization and spatio-temporal expression of *Hox* genes, which has been defined as "spatial" and "temporal" co-linearity (Figure 0.6A,B,C) (for review see Krumlauf, 1994; Kmita and Duboule, 2003). In simple words, this means that *Hox* genes located 3' within a given cluster are expressed earlier (temporal co-linearity) and with a more anterior boundary of expression (spatial colinearity) than 5' genes. Combinatorial *Hox* expression has been observed in all axial and paraxial tissues from the middle of the hindbrain to the tailbud, including paraxial and lateral plate mesoderm, spinal cord, neural crest, limbs, hindbrain segments, surface ectoderm, branchial arches, gut and gonadal tissues, thus suggesting the pivotal role of the Hox code in the A-P patterning of the whole embryo (Krumlauf, 1994).

One of the best examples of the role of the Hox code in the regulation of A-P patterning can be seen in the formation of the vertebral skeleton. Vertebrae derive from somites, which are blocks of paraxial mesoderm that form in a rostral-to caudal sequence between e8.0-e14.0 (Figure 0.6D) (Nagy et al., 2003). Somites then differentiate into dermatome, myotome and sclerotome, the later being the anlagen of the vertebrae, occipital bones and ribs. Each vertebrae is formed by the fusion of the caudal and rostral half of the sclerotome compartment of two adjacent somites in a process called re-segmentation (Dubrulle and Pourquie, 2004). This process generates the diversity of vertebrae that form the mouse vertebral skeleton (Figure 0.6E), organized from A to P order in 7 cervical (C1-C7), 13 thoracic (T1-T13), 6 lumbar (L1-L6), 3 or 4 sacral (S1-S4) and 31 caudal (Lohnes, 2003; Nagy et al., 2003). The morphology of the vertebrae largely differs between regions, and within a given region. For example, in the cervical region, C1 (or atlas) can be distinguished from C2 (axis), by the presence of a ventrally located tubercle known as the anterior arch

of the atlas (AAA) and a larger neural arch. C6 exhibits a characteristic TA see (Lohnes, 2003).

It is well demonstrated that, although morphologically indistinguishable one from another, each somite is poised with a particular Hox code, which determines its fate in accordance with its position along the A-P axis. The first somites that form are fated to become the most anterior pre-vertebrae by expressing 3' *Hox* genes, whereas the next forming somites acquire a more posterior identity by expressing a combination of 3' *Hox* plus the following more posterior 5' *Hox* genes. Disruption of this Hox code leads to vertebrae homeotic transformations, in which one particular vertebrae takes the identity of another. At the molecular level, homeotic transformations are the result of a change in somite identity resulting from a shift in the anterior expression boundary of *Hox* genes. For example, *Hoxd3* anterior boundary is at the level of somite 5 (see Lohnes, 2003). In *Hodx3*-null mouse mutants the expression domains of more anteriorly expressed *Hox* genes become posteriorly extended, which leads to anterior transformations of the first and second cervical vertebrae (C1 and C2) in basioccipital bones (Condie and Capecchi, 1993). On the basis of the phenotypes of loss-and gain-of function *Hox* mutants, it is generally accepted that loss or reduction of *Hox* gene causes anterior transformations whereas over-expression causes extension of these domains and posterior transformations (Lohnes, 2003). Interestingly, work from Olivier Pourquie's lab showed that *Hoxb* genes (i.e. *Hoxb1*, *Hoxb4*, *Hoxb7* and *Hoxb9*) are activated in a collinear fashion in epiblast cells that surround and ingress into the PS, and that this collinear activation of *Hoxb* genes control the flow of epiblast ingression for the correct establishment of the *Hox* code and the AP patterning of the paraxial mesoderm (Iimura and Pourquie, 2006).

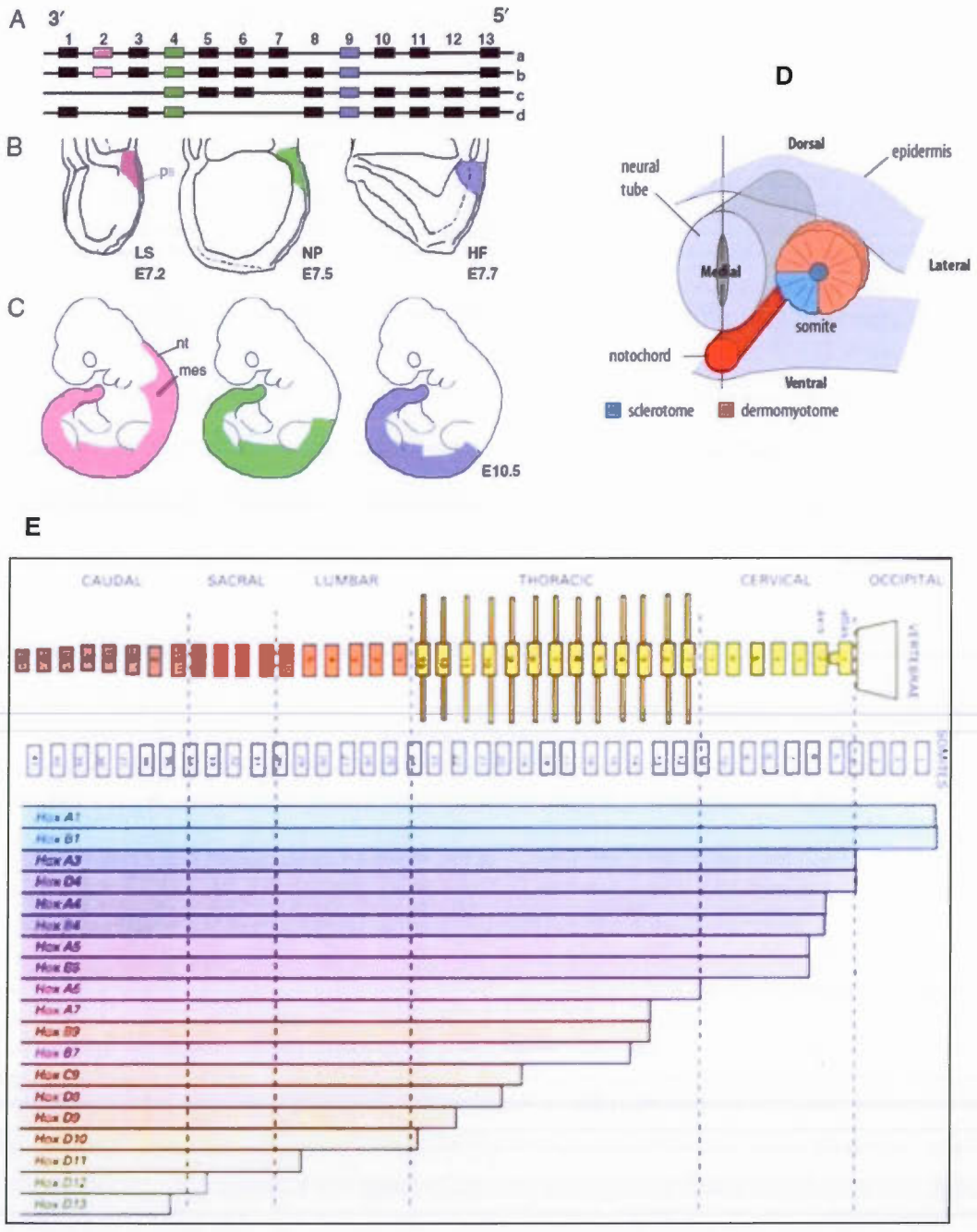


Figure 0.6 The murine *Hox* code and the control of anterior-posterior (A-P) patterning of vertebrae.

(A) Representation of the genomic arrangement of murine *Hox* genes. Taken from Deschamps and van Nes, 2005. The 39 *Hox* genes are distributed in four clusters in the

genome (*Hoxa* to *Hoxd*). *Hox* genes with the same number (1 to 13) are named paralogs. (B,C) Spatial and temporal colinearity of *Hox* genes illustrated for the expression of three paralogs groups: 2 (pink), 4 (green) and 9 (purple) at different developmental stages. Note that genes located 3' (e.g. *Hox2* paralogs) are expressed earlier and more anteriorly than 5' *Hox* genes (e.g. *Hox9* paralogs). Also note that for each gene, the expression boundary is more anterior in the neuroectoderm than in the mesoderm. nt: neural tube; mes, mesoderm; LS: late streak; NP: neural plate; HF: head fold. (D) Schematic representation of the somite and its compartments. (E) Relationship between the A-P location of somites, the A-P organisation of the vertebral skeleton and the Hox code. Adapted from Nagy et al., 2003.

0.4 Development of the nervous system

0.4.1 Neural induction

The first step in the formation of the nervous system is the specification of the neuroectoderm, a process known as neural induction. This process begins at gastrulation around mid-streak stage (e7.0) (Figure 0.7) and is spatiotemporally related to the establishment of the D-V axis. Cumulative data from studies in amphibians have led to the proposition of a model for explaining neural induction (Hemmati-Brivanlou and Melton, 1997); also see for review (Levine and Brivanlou, 2007; Munoz-Sanjuan and Brivanlou, 2002; Ozair et al., 2013). This model is known as the "the default model" and implies that inhibition of BMP signaling is the major event responsible for instructing ectoderm cells to become neural cells (Munoz-Sanjuan and Brivanlou, 2002). In absence of BMP inhibition, the ectoderm is "by default" instructed to become epidermis (Hemmati-Brivanlou and Melton, 1997). Some studies in frog and chick embryos suggest that neural induction also requires FGF signaling or alternative pathways (Ozair et al., 2013; Stern, 2005).

0.4.1.1 The default model: Lessons from *Xenopus*

Historically speaking, the first model for neural induction was proposed by Spemann and Mangold as early as 1924, based on their famous organizer grafting experiment performed in salamander gastrula embryo (Spemann, 1924). Briefly, grafting of the organizer tissue from a gastrula donor embryo into the ventral ectoderm region of

another early gastrula embryo induced the formation of dorsal tissues such as an ectopic nervous system at the expense of epidermis. According to this discovery, Spemann and Mangold proposed that the epidermis is the default fate of the ectoderm, and that positive signals emanating from the organizer induce the ectoderm to become neural tissue (Spemann, 1924). This was called the "epidermis by default model". Many years of research were dedicated to uncover the identity of these mysterious positive neural inductive signals, without results.

Around the 1990's decade, a number of studies performed with ectodermal explants of *Xenopus* gastrula embryos challenged the "epidermis by default model" of Spemann and Mangold and led to the proposition of a "neural default model" (Hemmati-Brivanlou and Melton, 1997). In 1989, it was reported that competent ectodermal cells of *Xenopus* (also known as animal cap), when explanted and cultured formed epidermis, but when dissociated (thus inhibiting cell-cell communication) formed neural tissue (Grunz and Tacke, 1989). This experiment was performed in the absence of mesoderm and endoderm tissue (organizer) or any signaling molecule and thus contradicted the idea that a positive signal was required for neural induction. Indeed, the observation strongly suggested that the neuroectoderm is the default state of the ectoderm and that a neural inhibitory signal was required to induce epidermis fate. The identity of this neural inhibitor and epidermal inducer, BMP signaling, was then clearly demonstrated by Brivanlou's group and others (reviewed in Levine and Brivanlou, 2007; Ozair et al., 2013). One of the most convincing work showed that exogenous administration of BMP4 to dissociated ectoderm cell explants represses neural identity and specified these cells to epidermis (Wilson and Hemmati-Brivanlou, 1995).

The mystery of the Spemann's organizer neural inducing activity was resolved in parallel to the understanding of its role in the formation of the D-V axis. The organizer-secreted factors chordin, noggin and follistatin, were shown to exhibit both

dorsalizing and neural inducing activity in *Xenopus* (Fainsod et al., 1997; Hemmati-Brivanlou et al., 1994; Lamb et al., 1993; Sasai et al., 1994; Smith and Harland, 1992). As described before, these factors act as BMP antagonists via direct binding and sequestering of BMP molecules in the extracellular space (Fainsod et al., 1997; Piccolo et al., 1996; Zimmerman et al., 1996). These discoveries strongly supported a default mechanism for neural induction in which the organizer secretes inhibitors of a neural inhibitor, rather than releasing a positive neural inductive signal for the ectoderm. This same mechanism, the BMP-chordin; noggin regulatory pathway has been shown to control allocation of ectoderm and mesoderm derivatives along the D-V axis (De Robertis and Kuroda, 2004), consistent with the dual role of the organizer in the control of neural induction and D-V patterning. It is important to point out that, although the blockage of BMP signaling by antagonist molecules is a canonical mechanism for neural induction, other antagonist-independent pathways have been shown to reinforce BMP inhibition, such as FGF and Wnt signaling (Munoz-Sanjuan and Brivanlou, 2002).

0.4.1.2 Neural induction in mouse

The neural default model works: Requirement of BMP signaling inhibition for neural induction

In the mouse, BMP signaling is active at the right time and space to control ectodermal cell fate decisions. Before gastrulation (e5.5), the distribution of BMP activity, (as detected by the levels of pSmad1,5,8) is observed in the visceral endoderm and also in a mosaic pattern in the epiblast (Di-Gregorio et al., 2007). At early gastrulation (e6.0-e6.5), the domain of BMP activity is restricted to the proximal epiblast and proximal visceral endoderm (de Sousa Lopes et al., 2003; Di-Gregorio et al., 2007; Hayashi et al., 2002). At mid-gastrula (e7.0), when neural induction begins, BMP activity is excluded from the prospective neural tissue that expresses the early neural marker *Otx2* (Ang et al., 1994). Indeed, BMP activity is

restricted to the posterior proximal epiblast region and only partially overlaps with the domain of canonical Wnt, FGF, Nodal and RA activity which are located more anteriorly in the PS at e7.0 (Figure 0.7A).

The understanding of the role of BMP signaling in mouse neuroectoderm specification did not go as fast as in *Xenopus*. Perhaps the most informative BMP loss-of-function mouse model is the epiblast-specific *Alk3* null mutant (Di-Gregorio et al., 2007). In this model the type I BMP receptor Alk3 was conditionally inactivated in the epiblast, by using a *Sox2*-Cre line. Interestingly, in the absence of BMP signal transduction almost the entire epiblast was prematurely differentiated to anterior neural tissue, at the expense of other tissues. This clearly revealed two central roles of BMP/Alk3 signaling in the mouse. First, BMP signaling is required to maintain epiblast pluripotency and second, inhibition of BMP signaling is essential for neural induction. Loss-of-function mutants involving inactivation of BMP antagonists Noggin and Chordin revealed functional overlap between these proteins in the control of neural specification and other developmental processes. Indeed, whereas single inactivation of Noggin or Chordin BMP antagonists in the mouse does not seem to affect neural specification, *noggin*$^{-/-}$ :: *chordin*$^{-/-}$ double mutants exhibit forebrain truncation but still form neural tissue (Bachiller et al., 2000). Importantly, these mutants exhibit malformed dorsal mesoderm and neuroectoderm structures that also denote defects in D-V patterning as well as defects in left-right asymmetry. This model without doubts demonstrated the pivotal role of BMP signaling in coordinating the formation of the three main axes.

Early neuroectoderm patterning: Posteriorizing signals come to play

An important topic regarding neural induction is how the neuroectoderm acquires an anterior (forebrain) or posterior (midbrain, hindbrain, spinal cord) character. The current view is based on the Nieuwkoop's "activation-transformation" model (Nieuwkoop, 1954). Nieuwkoop proposed that all the neural tissue is activated first

with an anterior character and is then transformed or posteriorized to form caudal neural tissue. One of the most convincing evidence that support the pre-specified anterior neural character of the embryo is the spatiotemporal expression pattern of *Otx2*, an early anterior neural marker. *Otx2* is a marker of forebrain and midbrain neural tissue (Ang et al., 1994). From pre-streak to early-streak stage (e6.0-6.5) *Otx2* is broadly expressed in the mouse epiblast (Ang et al., 1994). At the onset of neural induction by mid-streak stage (e7.0), *Otx2* expression is progressively excluded from the posterior region of the embryo and become juxtaposed to the domain of BMP, RA and Wnt activity (Fossat et al., 2012; Levine and Brivanlou, 2007) Figure 0.7A. Hence, the fact that the epiblast cells are initially marked by *Otx2* suggests that all cells of the embryo are initially fated to become the forebrain until signals instruct them to become the spinal cord, epidermis, mesoderm or endoderm.

It has been proposed that signals from FGF, Wnt and RA pathways are the Nieuwkoop's posteriorizing factors because: 1) their domain of activity is restricted to the posterior embryo (Garcia-Garcia and Anderson, 2003; Maretto et al., 2003; Rossant et al., 1991; Sun et al., 1999); and 2) the three signals can induce expression of early posterior neural markers. For example, stimulation of *Xenopus* gastrula animal caps with basic fibroblast growth factor (bFGF) induces the expression of the spinal cord neural marker *Hoxb9* (Lamb and Harland, 1995). As previously mentioned, studies in amphibians and chick have also shown a role for FGF signaling in neural induction (reviewed in Stern, 2005). It is not well understood whether FGF is required for neural induction in the mouse, since inhibition of FGF signaling in wild type and *Alk3*-null pre-gastrula embryo (described above), or in ES cells has no effect on neural specification (Di-Gregorio et al., 2007; Smukler et al., 2006). RA is a known inducer of neural differentiation in mES and P19 cell lines (Cai and Grabel, 2007; Jones-Villeneuve et al., 1983). However, mouse mutants for the RA synthesis enzyme Raldh2 – thus exhibiting defective RA signaling – do not fail to induce neural tissue (Molotkova et al., 2005). Further spinal cord neuronal differentiation is

severely impaired in *Raldh2* null mutants, thus suggesting that RA signaling is most likely implicated in posteriorization of the neurectoderm (Molotkova et al., 2005). In *Xenopus* animal caps assays RA represses anterior neural gene expression such as *Otx2* and induces expression of the posterior marker *Hoxb3* (Papalopulu and Kintner, 1996). Consistent with the results in *Xenopus*, addition of RA to mouse embryos represses *Otx2* expression and causes an anterior shift in the boundary of posterior genes such as the 3' *Hox2* genes (Ang et al., 1994; Conlon and Rossant, 1992). Canonical Wnt signaling appears to play a central, evolutionarily conserved role, in the A-P patterning of the neuroectoderm. The next section summarizes the evidences that support this current view.

Canonical Wnt posteriorizing signal

In all vertebrates the spatiotemporal domain and strength of Wnt activity must be fine-tuned during embryogenesis to ensure formation of the head (Arkell and Tam, 2012; Niehrs, 1999). In mice, as described above, the work of the DVE and AVE, that secrete powerful Wnt antagonist molecules exclude Wnt signaling from the prospective anterior region at pre-gastrula stage (Arnold and Robertson, 2009). This regionalization is then maintained by a complex regulatory network of Wnt antagonists, receptors and transcription factors, reviewed in (Arkell and Tam, 2012). At gastrulation, the expression of the Wnt antagonist Dkk1 is restricted to the anterior epiblast and opposes the domain of Wnt activity, denoted by the expression of the Wnt3 ligand in the posterior epiblast (Lewis et al., 2008) (Figure 8A). Whereas antagonists and ligands are regionalized, other members of the pathway, such as ß-catenin and the LRP6 co-receptor are widely expressed along the A-P axis (Arkell and Tam, 2012; Kelly et al., 2004; Zhao et al., 2014). This generates an anterior low to posterior high gradient of Wnt signaling in the epiblast at gastrulation that is conserved in all bilaterians (Figure 0.8A,B,C).

In *Xenopus* and zebrafish, two phases of canonical Wnt signaling have been described. In the early phase, at blastula stage, Wnt signaling mediates formation of the organizer (Pelegri and Maischein, 1998; Wessely et al., 2001). Thus, Wnt activity promotes neural induction in an indirect manner. In the late phase, at gastrulation, canonical Wnt signaling antagonizes the organizer, and inhibits both dorsalization and anterior neural induction. The best demonstration of the later can be seen in the phenotypes of Xwnt8 gain- and loss-of-function mutants (Christian and Moon, 1993; Hoppler et al., 1996). Overexpression of Xwnt8 at the late phase generates embryos without head (forebrain truncation) and notochord, but expanded posterior tissues (Christian and Moon, 1993). By contrast, late inhibition of Xwnt8 by expression of a wnt8 dominant negative protein generates embryos with big heads and notochord but small trunk tissues (Hoppler et al., 1996). Interestingly, concerted inhibition of BMP and Wnt signaling by co-injection of Dkk1 and a dominant negative BMP receptor generates embryos with multiple heads, thus demonstrating that BMP and Wnt signaling pathways act synergistically in the inhibition of head induction (Glinka et al., 1997) (Figure 0.8D). In *Xenopus* animal cap assays, activation of Wnt signaling in cooperation with BMP inhibition represses anterior neural gene expression and concomitantly leads to activation of posterior neural markers (McGrew et al., 1995).

Several loss- and gain-of- Wnt signaling mouse models corroborate the results in *Xenopus* and zebrafish and thus suggest an evolutionarily conserved, posteriorizing role for canonical Wnt signaling. Ectopic activation of the Wnt-ßcatenin pathway by expression of the chick Wnt8 ligand in transgenic mouse embryos or null mutation of the major Wnt antagonist Dkk1 leads to forebrain truncation and expansion of midbrain and hindbrain tissue (Lewis et al., 2008; Mukhopadhyay et al., 2001; Popperl et al., 1997), (see also Figure 0.8B). On the other hand, decrease in canonical Wnt signaling by mutation of the Wnt3a ligand or the co-receptors Lrp5 and Lrp6 leads to expansion of neural tissue at the expense of posterior structures (Kelly et al., 2004; Yoshikawa et al., 1997). As far as we know, the effect of repressing both BMP

and Wnt signaling *in vivo* has not been reported, nor yet described multi-head mouse embryos. In both *Xenopus* and mouse, BMP and Wnt signaling synergize in repression of anterior neural induction. Indeed, simultaneous loss-of-function of the inhibitors Noggin and Dkk1 *in vivo* leads to progressively more severe anterior truncation phenotypes (del Barco Barrantes et al., 2003).

In support to Nieuwkoop's activation-transformation model, all this evidence has lead to the proposition of a model in which inhibition of both BMP and canonical Wnt signaling is required for a first step of anterior neuroectoderm induction, whereas inhibition of BMP and activation of Wnt signaling is required for further posteriorization of the neuroectoderm (Niehrs, 1999).

Integrating the information: Spatiotemporal control of neural induction in the mouse

Before gastrulation at pre-streak stage (e6.0-e6.5), the epiblast appears to exist as a pre-specified anterior neural state as evidenced by the wide expression of the anterior neural marker *Otx2* (Ang et al., 1994; Levine and Brivanlou, 2007) (Figure 0.7B). The location of the AVE delimits the prospective anterior pole of the embryo, where the anterior neuroectoderm is specified by default (Arnold and Robertson, 2009). The inhibitory action of AVE protects this region from the action of BMP, Nodal and Wnt signals, which get restricted to the other side of the embryo and induce formation of the PS at e6.5.

Formation of the PS marks the onset of gastrulation and sets the A-P axis. Epiblast cells that receive BMP, Wnt and Nodal signals ingress through the posterior region of the PS and give rise to mesoderm and definitive endoderm lineages. Epiblast cells located at the most anterior end of the PS at early-gastrula stage form the GO signaling center (Kinder et al., 2001). The GO formed by EGO and MGO is a dynamic cell population that secretes powerful antagonists of BMP, Nodal and Wnt signaling. Epiblast cells under the action of these inhibitory molecules do not receive

any signaling and form anterior neuroectoderm by default at mid-streak stage (e7.0). As gastrulation progress concomitantly with embryo elongation, the anterior neural tissue moves away from the GO. The anterior neuroectoderm is maintained by the action first of the AVE and then by the AME derivatives of the GO that ingress through the PS and localize at the anterior midline of the embryo, replacing the AVE (Figure 0.7B) (Bachiller et al., 2000; Beddington and Robertson, 1999; Yamanaka et al., 2007; Yang and Klingensmith, 2006).

At late PS stage (e7.5) the GO has become the node (Arnold and Robertson, 2009; Yang and Klingensmith, 2006). The node is thus involved in the specification of posterior neural tissue fated to become midbrain/hindbrain/spinal cord (Figure 0.7B) (Beddington and Robertson, 1999; Levine and Brivanlou, 2007). Consistent with the Nieuwkoop's activation-transformation model, the node induces neural tissue with an anterior character that is subsequently posteriorized by the action of canonical Wnt, FGF and RA signals in close proximity to this organizer (Levine and Brivanlou, 2007). As the specified posterior neuroectoderm moves away from the node concomitantly with embryo elongation, the derivatives of the node including the posterior notochord lie at the ventral midline and continue to secrete BMP antagonists (Kinder et al., 2001; Yamanaka et al., 2007; Yang and Klingensmith, 2006).

At headfold stage (e8.5), the neuroectoderm is already patterned into forebrain marked by expression of *Otx2*, midbrain marked by expression of *En2*, hindbrain marked by *Krox20* expression and spinal cord domains marked by expression of *Hoxb1* (Ang et al., 1994; Bachiller et al., 2000). From this stage until ~e13.5 the node and the PS have regressed caudally to a posterior stem cell zone where new epiblast progenitors are constantly generated. In the open posterior neural plate (PNP) region, the inhibitory action of the node and the tail notochord restrict BMP signaling to the lateral ectoderm generating a mediolateral gradient of BMP signaling (Yang and Klingensmith, 2006) (Figure 0.5), whereas Wnt activity is broadly distributed in the

PNP and other tissues (Maretto et al., 2003; Yang and Klingensmith, 2006) (Figure 0.8D). The notochord also secretes Shh morphogen that together with BMP signaling plays critical roles in further steps of neural patterning.

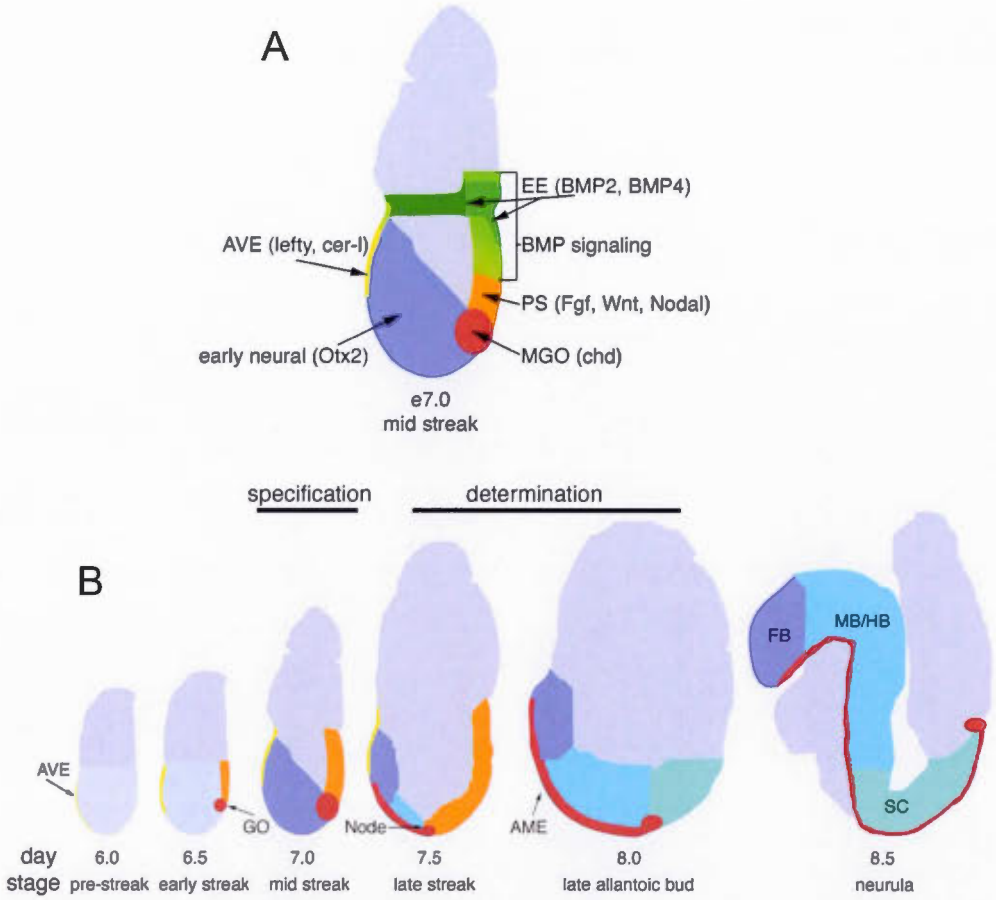


Figure 0.7 Neural induction in the mouse embryo.

Taken from Levine and Brivanlou, 2007 (A) Distribution of organizing centers and signaling molecules during neural induction at mid-streak stage e7.0. AVE: anterior visceral endoderm; PS: primitive streak; MGO: mid-gastrula organizer; chd: chordin; cer-1: Cerberus; EE: extra-embryonic region. (B) Neural induction and neuroectoderm patterning from e6.0-e8.5. From e6.0 to e6.5 the epiblast exist in an anterior pre-neural state with the AVE underlying the anterior epiblast and the anterior neural marker *Otx2* being broadly expressed in the epiblast (blue). The onset of gastrulation (e6.5) is denoted by the formation of the PS and the early gastrula organizer (GO) in the most anterior part of the PS. From the onset of neural induction onwards (e7.0), *Otx2* expression becomes progressively excluded from the posterior region and juxtaposes the domain of signaling activity (shown in A). The anterior neuroectoderm is specified by default from epiblast cells that do not ingress the PS and do not receive signaling. Its character is maintained by the anterior mesendoderm (AME). By

late streak stage (e7.5), the node induces neural tissue with an anterior character that is subsequently posteriorized to become midbrain (MB) /hindbrain (HB) (light blue)/spinal cord (SC) (green) by the action of canonical Wnt, FGF and RA signals.

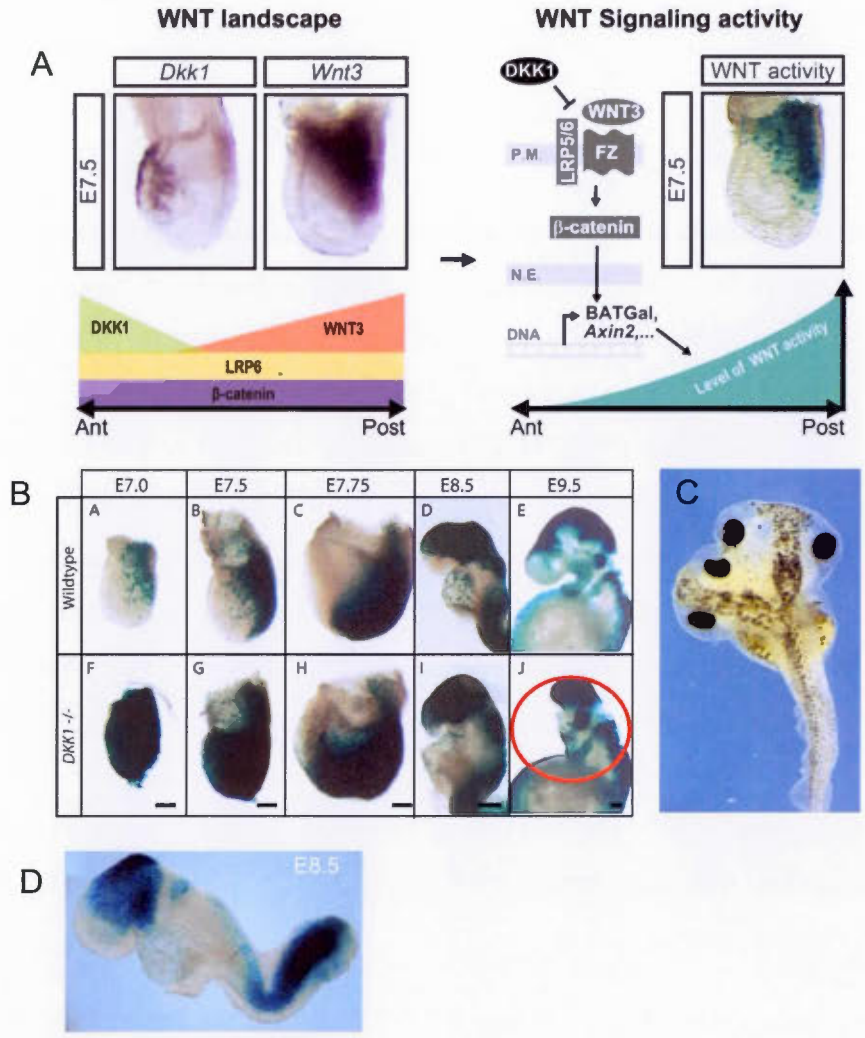


Figure 0.8 Posteriorizing canonical Wnt signaling.

(A) Representation of the distribution of canonical Wnt signaling components and Wnt antagonists in the mouse embryo. Taken from Fossat et al., 2012. (B) Distribution of Wnt signaling activity from e7.0-9.5 as assessed by expression of the BAT-Gal reporter in wild-type and *Dkk1*-null embryos. Source of the figure: Lewis et al., 2008. The BAT-Gal reporter drives *lacZ* expression under the control of multimerized LEF/TCF binding sites. See the expansion in the domain of Wnt activity and the consequent forebrain truncation in *Dkk1* mutants. (C) Formation of a secondary head in a tadpole-stage *Xenopus* embryo co-injected with mRNA encoding Wnt and BMP inhibitors. Taken from Niehrs, 1999. (D) Profile of

activated Wnt signaling in an e8.5 mouse embryo. Note the broad domain of Wnt activity in the posterior neural plate.

0.4.2 Neurulation

0.4.2.1 Primary neurulation

Following neural induction at e8.5 (~6-somite stage), the neural plate begins to roll and form the neural tube (NT), the precursor of the central nervous system (CNS). This apparently simple process is known as primary neurulation and requires complex morphogenetic movements as well as precise control of apoptosis and cell proliferation (Copp et al., 2003). Before neurulation the neural plate has the shape of a disc, and must 1) elongate and become narrower, 2) fold and elevate and 3) fuse at the midline, in order to form the NT (Figure 0.9A). Perhaps because of the complexity of this process, the NT forms not at once, but in a discontinuous manner along the A-P axis in mammals. In mouse and human species fusion initiates at three spatiotemporally separated sites which are commonly known as closure 1, 2 and 3; then proceeds in a zipper like-manner from these sites to form the brain and most of the spinal cord (Figure 0.9B). However, it is important to point out that there is controversy regarding whether there is a closure 2 event in human (Figure 0.10) (reviewed in Copp and Greene, 2013).

Before primary neurulation begins, at late PS (e7.5) the architecture of the neural plate and the whole embryo is remodelled by a morphogenetic movement known as convergent extension (CE). In this process, neural and axial mesoderm cells located anterior to the node move to the midline and intercalate (Figure 0.9A) (Copp et al., 2003; Ybot-Gonzalez et al., 2007b). Cell reorganization leads to embryo elongation and thus to lengthening and narrowing of the neural plate. Midline convergence of axial mesoderm cells leads to formation of the notochord, which signals to the overlying neuroectoderm and induce neural plate bending at its median hinge point

(MHP) (Figure 0.9C). This leads to formation and elevation of the neural folds (Copp et al., 2003).

Studies in several models suggest important roles for the underlying mesoderm in the coordination of NT closure. In the mouse, proliferation of the cranial mesenchyme is essential for cranial neurulation, since inactivation of the cranial mesenchyme marker *Twist* leads to cranial NTDs (Chen and Behringer, 1995). It is, however, important to point out that spinal neural tube closure was not affected by removal of the paraxial mesoderm, thus suggesting a different contribution of the mesoderm to NT closure along the AP axis in mouse. Studies in other models such as zebrafish and *Xenopus* suggest that the mesoderm is required for the coordination of movements of neural plate cells during neurulation (Araya et al., 2014).

Around the 6-somite stage (e8.5), at the hindbrain-spinal cord boundary, the V-shaped neural folds become close enough to appose and fuse, marking the first site of NT closure (known as closure 1) (Figure 0.9A). Approximately 12 hours after initiation of closure 1, closure 2 and 3 are found at the forebrain/midbrain boundary and at the rostral end of the forebrain, respectively (Figure 0.9B). NT closure proceeds bi-directionally in a zipper-like manner from closure sites 1 and 2, and in a caudal direction from closure site 3. Cranial neurulation is completed a few hours after closure 2 and 3. On the other hand, dorsolateral bending and closure of the posterior neuropore is assisted by the formation of paired dorsolateral hinge points (DLHPs) at the site of attachment of the PNP with the surface ectoderm in intermediate and lower regions of the spinal cord (Ybot-Gonzalez et al., 2007a) (Figure 0.9C). Spinal cord neurulation finishes with closure of the posterior neuropore at the upper sacral level around e10 (~30-somite stage) (Copp and Greene, 2013; Copp et al., 2003).

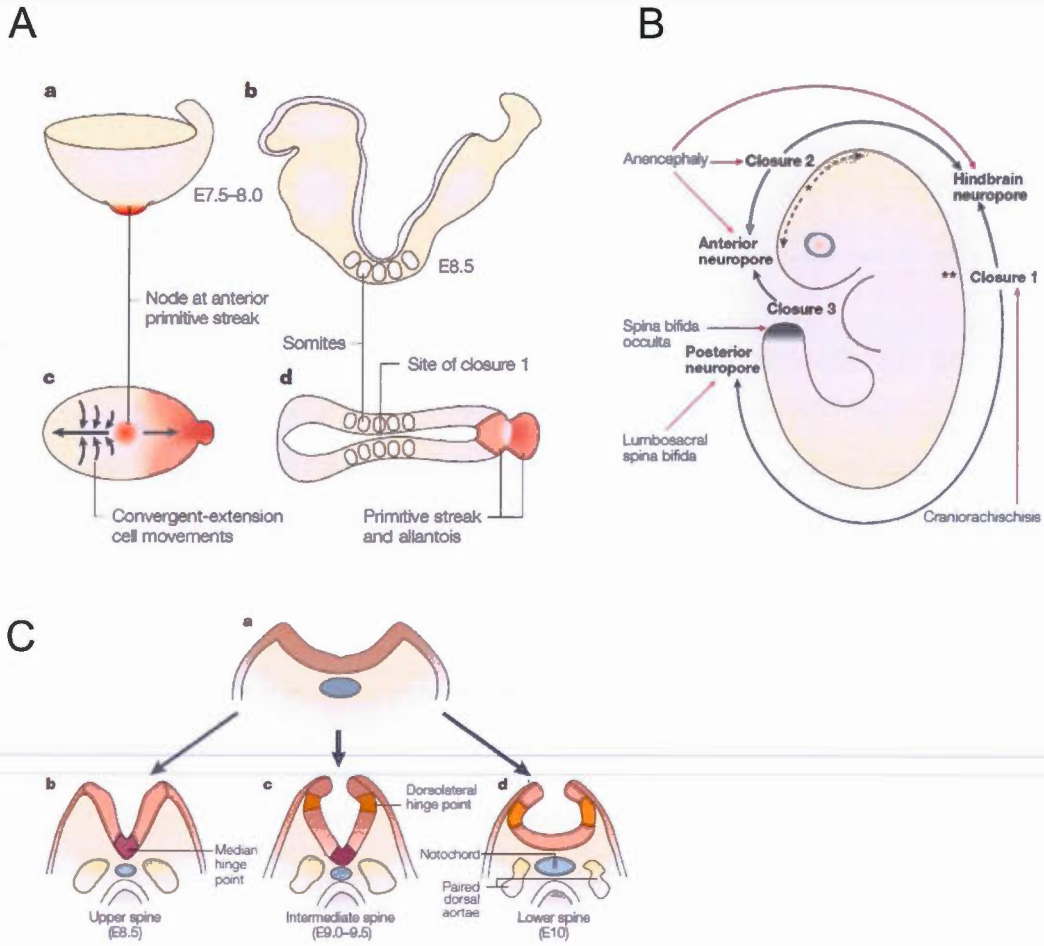


Figure 0.9 Spatio-temporal sequence of neurulation events in the mouse embryo.

Taken from Copp et al., 2003. (A) Schematic representation of the shaping of the neural plate by the morphogenetic movement of convergent extension that ensures initiation of neural tube closure at e8.5 at the hindbrain-spinal cord boundary (site of closure 1). (B) Rostro-caudal sequence of neurulation events and NT defects that arise when this process fails. After initiation of closure 1, two other *de novo* closure sites known as closure 2 (at the forebrain/midbrain boundary) and closure 3 at the rostral end of the forebrain ensure completion of neurulation. NT closure spreads bi-directionally in a zipper-like manner from closure sites 1 and 2, and in a caudal direction from closure site 3. (C) Changes in morphology of the neural plate and location of bending points: the median hinge point and dorsolateral hinge points at different levels of the A-P axis that facilitate neurulation.

0.4.2.2 Secondary neurulation

Following primary neurulation, the rest of the spinal cord, at lower sacral and coccyx regions (mouse and human) and tail in mouse, forms by secondary neurulation. At the difference of the primary NT which originates from the neural plate, the secondary NT originates from tailbud mesenchymal cells (Schoenwolf, 1984). From e9.5-10, these cells condense forming a rod like structure known as the medullary cord. The hollow secondary NT forms when mesenchymal cells of the medulary cord are transformed into a neuroepithelium and arrange radially around a central lumen formed by cavitation. From e11-12, the rest of the secondary neural tube forms from a medullary plate rather than a cord, and by addition of tailbud cells to the edges of the plate. Maybe because of its poor contribution to the CNS, the mechanisms that control secondary neurulation are not well understood. The neural cell-adhesion molecule (N-CAM), the secreted protein midkine and BMP signaling are proposed as candidates for mediating the mesenchymal to epithelial transition of secondary neurulation (Lowery and Sive, 2004).

0.4.2.3 "Open" and "Closed" Neural Tube Defects

Failure of neurulation results in a group of serious congenital malformations known as neural tube defects (NTDs). After congenital heart defects, NTDs are the second most frequent group of human birth anomalies, affecting 1 per 1000 pregnancies (Mitchell, 2005). NTDs are multifactorial and result from the additive contribution, and interaction, of genetic and environmental risk factors (Greene and Copp, 2014).

Failure of primary neurulation results in "open" NTDs, in which the NT lumen is exposed to the outside environment and the function of the nervous system is severely compromised (Figure 0.10). The major open NTDs are anencephaly, craniorachischisis and open spina bifida (Copp and Greene, 2013). In both anencephaly and craniorachischisis the brain is exposed, leading to embryonic or perinatal lethality. Babies with open spina bifida (also known as myelomeningocele,

spina bifida aperta) survive, but frequently suffer from inability to walk, incontinence and lack of sensitivity in the affected region of the spinal cord (Copp and Greene, 2013). Failure of secondary neurulation leads to "closed" NTDs in which the spinal cord forms but fails to detach from tailbud surrounding tissues. This class of NTDs may present in asymptomatic forms such as spina bifida oculta, also known as skin-covered spinal 'dysraphism', or as severe spinal cord tethering affecting lower limb motility (Copp et al., 2003).

Open NTDs occur when closure 1, 2 or 3 fail or from defects in the zippering of the neuropores (Figure 0.9C and 0.10). Craniorachischisis is associated with failure of closure 1, and is the most sèvere NTD in which the NT remains open from the midbrain down the whole spinal cord region (Copp et al., 2003). Exencephaly is commonly associated with failure in closure 2 or 3, and lead to anencephaly by late gestation, characterized by absence of the skull vault and brain tissue destruction. Open spina bifida occurs when the zippering after closure 1 fails to complete along the spinal cord region and the posterior neuropore remains open (Copp et al., 2003). The severity of the spina bifida lesion can vary depending on the developmental stage at which closure stops. For example, in the *Zic2 Kumba* mouse mutant, posterior neuropore fails to close as early as the 14-16-somite stage due to defects in formation of the DLHP, which leads to a large spina bifida beginning at the thoracic level (Elms et al., 2003).

A study in 2010 reported more than 240 mouse mutants with NTDs, with more than 200 genes implicated in the control of neurulation (Harris and Juriloff, 2010). On the basis of this study exencephaly/anencephaly is the most frequent NTD, followed by spina bifida which can present alone or with exencephaly. Craniorachischisis is the less common NTD, and arrives mainly by defects in function of planar cell polarity (PCP) pathway genes (13/14 genes are PCP related). Several PCP mutants exhibit craniorachischisis in high penetrance (95-100%), as is the case of loop-tail (*Vangl2*),

Celsr1 as well as double compound mutants for Dvl (*Dvl1/Dvl2; Dvl2/Dvl3)* or Fzd (*Fzd3/ Fzd6*) (Merello et al., 2014). On the other hand, it has been shown that genetic background considerably affects the severity, type and penetrance of NTDs (Harris and Juriloff, 2010).

In spite of the high number of mouse NTD genes, only few human orthologues have been identified as predisposing factors for NTDs (Copp and Greene, 2013). Almost all human mutations affect genes implicated in the control of CE movements via the PCP pathway (11 PCP genes identified as predisposing factors for human NTD so far) (Merello et al., 2014). This demonstrates the central role of PCP signaling in the regulation of neurulation. A few non-PCP related genes such as the dorsal NT marker paired box 3 transcription factor (PAX3) and the posterior mesoderm marker T (Brachyury) have also been associated with spina bifida in humans (Agopian et al., 2013).

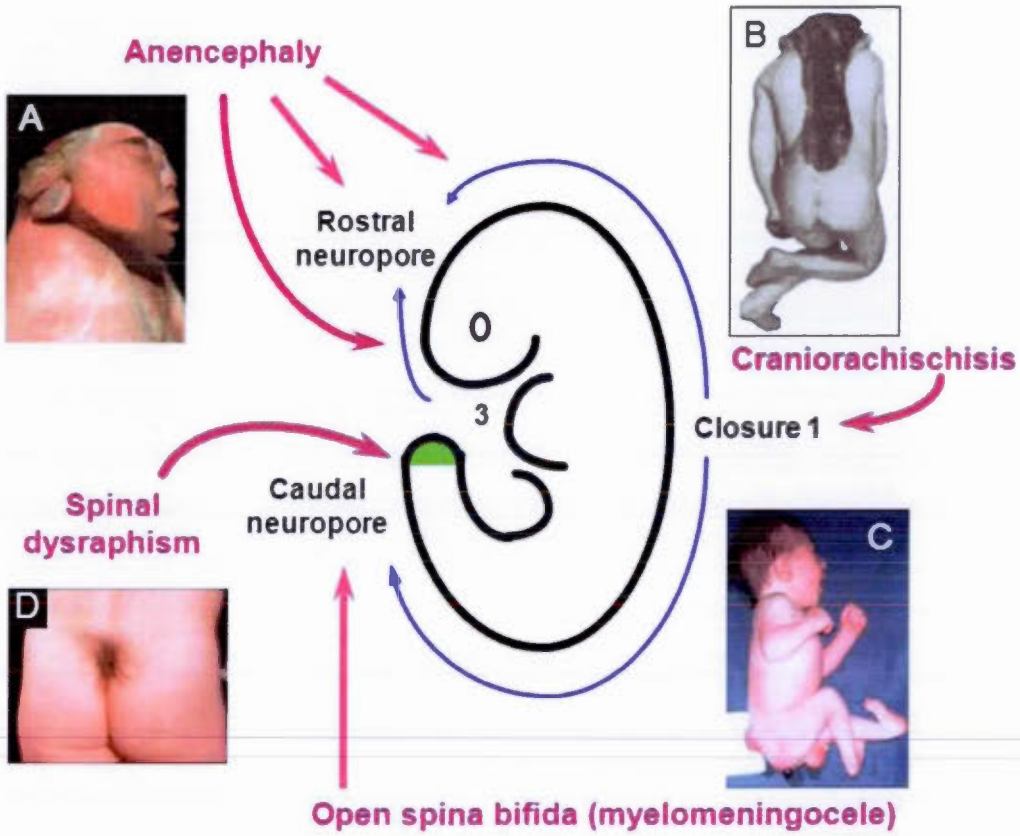


Figure 0.10 Neurulation in the human embryo and neural tube defects resulting from disturbed primary and secondary neurulation.

Source of the figure: Copp and Greene, 2013.

0.4.2.4 Molecular and cellular control of primary neurulation

On the basis of the more than 240 mouse genetic models of NTDs, a considerable number of molecular and cellular events have been implicated in the control of neurulation, demonstrating the complexity of this developmental process. Regulators of signaling from canonical and non-canonical (PCP) Wnt, Shh, BMP and RA pathways, folate one-carbon metabolic pathway, chromatin modification events, as well as cellular processes such as cell proliferation, survival and motility are all essential for NT closure (Copp and Greene, 2013; Greene et al., 2009; Zhao et al.,

2014). It is still not well understood how some of these regulators operate in the control of neurulation, as is the case, for example, of RA signaling (Copp and Greene, 2013). Analysis of the mouse NTDs models has also revealed that the mechanisms that control NT closure differ along the A-P axis (Copp et al., 2003). This is well exemplified by the differences in regulation of cranial and spinal cord neurulation. Indeed, it has been shown that cranial NT closure highly depends on the proliferation and expansion of the mesenchyme, contraction of sub-apical actin cytoskeleton, and emigration of the neural crest population. However, perturbation of any of these processes appears to not affect spinal cord neurulation (Copp and Greene, 2013; Copp et al., 2003). A description of the main signaling pathways and mechanisms controlling spinal cord neurulation follows.

PCP signaling is required for initiation of spinal cord neurulation

The first key event in spinal cord neurulation is the PCP-dependent shaping of the neural plate by convergent (narrowing) extension (lengthening). Perturbation of PCP signaling, as seen in loss-of function mouse mutants for several PCP genes (*Vangl2 (Lp mutant)*, *Celsr1*, *Dvl*, *Ptk7*), prevents CE movements and generally leads to failure in closure 1 and craniorachischisis (Copp et al., 2003; Kibar et al., 2001; Lu et al., 2004; Merello et al., 2014; Savory et al., 2011a). PCP mutant embryos are characterized by a short and wide body axis, including a broad neural plate whose neural folds are too far apart to close (Copp and Greene, 2013). How exactly PCP signaling regulates CE is not well understood. Some studies propose that PCP activation polarizes actin-rich lamellipodia to the medial and lateral ends of neural and mesodermal cells, thus contributing to elongation and alignment of cells to the mediolateral axis. Mediolateral cell intercalation then occurs when cell elongation leads to formation of lamelliform protusions that confer cell-cell and cell-matrix traction (Copp and Greene, 2013; Keller et al., 2008). More recently, live imaging analysis of neural plate CE in *Ptk7* and *Vangl2* PCP mouse mutants demonstrated a role for PCP signaling in the control of mediolateral neural cell intercalation by two

different mechanisms involving apical boundary rearrangement and polarized basolateral protrusive activity (Williams et al., 2014). This work showed that craniorachischisis in *Ptk7* mutant is caused by an impaired neural tissue polarization and apical constriction. The role of mesoderm CE was not studied (Williams et al., 2014).

A regulatory network comprised by BMP-noggin, Shh signaling and Zic2 control bending of the neural folds at spinal level

As described above, elevation and apposition of the neural folds during neurulation is facilitated by formation of two bending points, the MHP at the midline of the neural plate and a pair of DLHPs at the dorso-lateral region of the neural folds (Copp et al., 2003) (Figure 0.9C). Neuroepithelial cells at the bending point acquire a wedge shape (wider basally than apically). Although regulation of cell cycle has been proposed as a putative mechanism, the mechanisms that control these cell shape changes are not well understood (Greene and Copp, 2014). MHP and DLHP are differentially formed during the rostro-caudal wave of spinal cord neurulation. MHP is the only bending point at upper levels of the forming spinal cord from e8.5-9.0. DLHPs forms at intermediate levels of the spinal cord and coexists with the MHP from e9.0 to e9.5. As the wave of neurulation proceeds caudally, DLHP is the only bending point at lower levels of the forming spinal cord from e9.75 to e10.5 (Copp et al., 2003). Formation of the MHP requires strong levels of Shh signaling coming from the underlying notochord. It has been proposed that the MHP is dispensable for spinal cord neurulation, as this process completes successfully in *Shh*-null mutants which do not have a floor plate and thus a MHP (Chiang et al., 1996). However, formation of the DLHPs and thus dorsolateral bending of the neural plate in intermediate and lower regions of the spinal cord, is essential for NT closure (Greene and Copp, 2014; Ybot-Gonzalez et al., 2007a).

DLHPs formation is controlled by the interplay between dorso-ventral cues from BMP and Shh signaling – which inhibit DLHPs at upper spinal cord regions – as well as the BMP-antagonist Noggin and the zinc finger transcription factor Zic2 – which are both necessary for DLHP formation in the intermediate and lower spinal cord (Ybot-Gonzalez et al., 2007a). In both *Shh*- and *BMP2*-null mutants DLHPs are ectopically formed at anterior levels of the body axis (Ybot-Gonzalez et al., 2002; Ybot-Gonzalez et al., 2007a). Both *Noggin* (*Nog*$^{-/-}$) and *Zic2* loss-of-function (Zic2$^{Ku/Ku}$) mouse mutants lack DLHPs and exhibit severe spina bifida (Elms et al., 2003; McMahon et al., 1998). In support of a major role for BMP inhibition in several neurulation-related processes, Andrew J. Copps' lab proposed in 2007 that Noggin-mediated antagonism of BMP signaling is the event that induces formation of DLHPs and thus dorsolateral bending of the neural plate at intermediate and low spinal cord levels (Ybot-Gonzalez et al., 2007a). This work proposed a mechanism in which at upper spinal levels Shh signaling secreted from the notochord is strong and inhibits Noggin expression, thus leading BMP signaling to inhibit DLHP formation at this level. At lower levels of the spinal cord Shh signaling strength decreases, resulting in Noggin expression and inhibition of BMP signaling thereby leading to formation of DLHP and dorsolateral bending of the neural plate. Importantly, Zic2 is involved in this mechanism by being necessary for Noggin expression at caudal levels, as expression of this BMP antagonist is absent in *Zic2*$^{Ku/Ku}$ mutants (Ybot-Gonzalez et al., 2007a). Whether Zic2 fulfills this role indirectly via inhibition of Shh signaling, or directly via activation of *Noggin* expression is not well understood. Interestingly, studies in zebrafish have also revealled that the *zic2a/zic5* gene pair acts downstream of canonical Wnt signaling to control DLHP formation and cranial neural fold bending via modulation of the actomyosin cytoskeleton and junction integrity (Nyholm et al., 2009).

Canonical Wnt signaling and spinal cord development

Several downstream targets of Wnt-ßcatenin signaling such as the genes encoding Cdx, Pax3 and Zic transcription factors have been involved in the control of mouse spinal cord neurulation, thus suggesting a pivotal role for the canonical arm of Wnt signaling in the organization of this process (Epstein et al., 1991; Inoue et al., 2004; Savory et al., 2011a). A direct requirement for Wnt-ßcatenin signaling in posterior neuropore closure was recently demonstrated by using a conditional mutagenesis approach, in which ß-catenin was inactivated in the dorsal NT by using a *Pax3*Cre line (Zhao et al., 2014). ß-catenin conditional KO mutants exhibited open spina bifida and reduced expression of *Cdx2/4*, *Pax3* and *Msx1* genes (Zhao et al., 2014). Expression of *Pax3* rescued the spina bifida phenotype thus demonstrating that canonical Wnt acts via regulation of *Pax3* expression in the control of posterior NT closure. However, the mechanism by which Wnt-ß-catenin/*Pax3* control this process still remains unknown. These regulators appear to act on the last step of neurulation involving fusion of the apices of posterior neural folds, probably by regulating cell proliferation, adhesion and apoptosis (Copp et al., 2003). It is important to point out that, given the implication of ß-catenin in cell-cell adhesion; it is likely that defects in ß-catenin conditional KO mutants be associated with defective cell adhesion rather than misregulated canonical Wnt signaling.

0.5 Getting to the crest: the neural crest

0.5.1 Origin and evolution

Discovered about 150 years ago by Wilhem His (His, 1868), the neural crest (NC) is a transient population of multipotent and migratory cells that arise at a precise space and time in the vertebrate embryo. This amazing population received its named based on its location at the "crest" of the forming NT. Progenitors of the NC are initially induced at the neural plate border, which is the region at the junction between the neural and non-neural ectoderm (Figure 0.11A). In the mouse, NC induction is

evident at e8.5 with the onset of *Pax3* expression around the hindbrain/spinal-cord boundary, and continues at the lateral borders of the posterior neural plate concomitantly with axial elongation and generation of new neural cells from the posterior growth zone. Along with neural plate folding and formation of the NT, NC progenitors are specified into *bona fide* NCC and reside at the dorsal part of the NT. NCC experience an epithelial-to-mesenchymal transition leading to delamination from the NT, colonization of the embryo via defined migratory pathways and differentiation into many derivatives (Bronner and LeDouarin, 2012).

The NC has been called "the fourth germ layer" because of its enormous contribution to embryonic tissues (Hall, 2000). NCC give rise to multiple structures and mature cell types (Figure 0.11B), such as chondrocytes, osteocytes and fibroblasts of the craniofacial skeleton, neurons and glia of the peripheral and enteric nervous system, the thymus and organs derived from the pharyngeal region, smooth muscle cells, cardiac outflow tract septum, dorsal root ganglia (DRG), sympathetic chain ganglia, adrenal chromaffin cells, as well as melanocytes (pigment cells) (Bronner and LeDouarin, 2012; Green and Bronner, 2013).

The NC population is unique to vertebrates and its emergence has been proposed to be a key event in the transition from invertebrates to vertebrates and evolution of the new vertebrate head (Gans and Northcutt, 1983; Jandzik et al., 2014). Indeed, the main morphological differences that distinguish vertebrates from other chordates reside in the head and derive from the NC. The best example of this is the presence of jaws, which permitted ancestral vertebrates to shift from filter feeding to active predation. A particular interest has been put in the last years to the understanding of the origin and evolution of the NC by analysis of the expression of key NC regulatory genes in urochordates (Abitua et al., 2012; Jeffery, 2006; Jeffery et al., 2008; Jeffery et al., 2004), in the basal chordate amphioxus (Jandzik et al., 2014; Yu et al., 2008) as well as in the jawless vertebrate lamprey (Nikitina and Bronner-Fraser, 2009; Sauka-

Spengler et al., 2007). Based on many of these studies, Marianne Bronner's group has proposed that the evolution of the NC was facilitated by gene duplication and co-option of mesoderm genes into the neural plate border (Green and Bronner, 2013). However, studies performed in urochordate ascidian species, (which are considered the true sister group to vertebrates), challenge the hypothesis of NC origin by gene duplication (Abitua et al., 2012; Jeffery, 2006; Jeffery et al., 2008; Jeffery et al., 2004). Indeed, William R. Jeffery et al. identified migratory NC like cells (NCLC) in diverse ascidian species that express the neural crest marker HNK-1 and differentiate into body pigment cells (Jeffery, 2006; Jeffery et al., 2004). Cleavage arrest studies in the ascidian urochordate *Ciona intestinalis* showed that NCLC derive from the A7.6 lineage, which gives rise to a type of migratory mesenchymal cells known as trunk lateral cells (TLC) (Jeffery et al., 2008). It was shown that the A7.6/TLC lineage expresses a number of orthologues of the vertebrate NC regulatory genes, although it does not arises at the neural plate border and lack expression of neural plate border specifiers (Jeffery et al., 2008). Then after, Michael Levine's group proposed that the bilateral a9.49 pigment cell lineage of *Ciona* represents a rudimentary NC, which: 1) arises at the neural plate border; 2) expresses a number of orthologues of the vertebrate neural plate border specifier genes and NC specifiers and 3) can be reprogrammed into migrating "NC-like cells" by ectopic expression of a homolog of the key mesenchymal determinant gene Twist (Abitua et al., 2012). According to this result, Levine's group claimed that pigment cells or melanocytes are the ancestral precursors of NCC and that vertebrates segregated from basal chordates by co-option of mesenchyme determinants into the neural plate ectoderm (Abitua et al., 2012). However, some questions remain unanswered about this model and the origin of NCC is still controversial (reviewed in Green and Bronner, 2013).

0.5.2 NC populations across the A-P axis

NCC fate and migratory pathways are determined by their position along the A-P and D-V axis, as well as by the interaction with extrinsic signals (Bronner and LeDouarin, 2012; Gammill et al., 2006; Pavan and Raible, 2012). According to their A-P axis location, premigratory NCC are subdivided into cranial, cardiac, vagal, trunk and sacral populations. The cranial population has been extensively studied because of its contribution to the craniofacial skeleton. These cells also contribute to sensory neurons and glia of the face, thymus and other organs in the head. Most of the vagal population colonize the gut and give rise to the enteric nervous system, whereas a subgroup known as the cardiac NC colonize the heart and contributes to the outflow tract. Trunk neural crest cells give rise to DRG, sympathetic ganglia, Schwann cells, adrenal chromafin cells and to the pigment producing cells of the skin (melanocytes). The sacral population contributes to the enteric nervous system, but only in the most distal region of the gut (Bronner and LeDouarin, 2012).

0.5.3 Differences in NCC behaviour along the A-P axis and between species

In general, NCC delamination follows a rostro-caudal sequence. However, the timing of delamination largely differ in accordance with the level of the A-P axis and between species (Theveneau and Mayor, 2012). In mouse and *Xenopus*, cranial NCC delaminate all at once, before closure of the cranial neuropore, whereas in birds cranial NCC delaminate concomitantly with NT closure. In contrast to the cranial population, in every vertebrate species delamination of the trunk NCC population takes place progressively in a rostro-caudal fashion and several hours after completion of NT closure (Erickson and Weston, 1983). However, the timing of trunk NCC delamination may vary depending again on the position along the A-P axis. Interestingly, analysis of mouse mutants exhibiting spina bifida, as is the case of the *Pax3* splotch mutant or compound splotch/curly tail mice mutants, suggest that

NCC delamination is functionally dissociated from NT closure, since NCC normally emigrate from open NTs (Estibeiro et al., 1993; Franz, 1992).

The migratory pathways of cranial NCC are largely conserved in vertebrates. However, the migratory behavior of trunk NCC populations differs largely from that of the other populations and between species (Bronner and LeDouarin, 2012; Theveneau and Mayor, 2012). Indeed, trunk NCC migrate either ventrally to give rise to neural derivatives, or dorsolaterally to give rise to melanocytes. Species-specific differences are observed in the way trunk NCC migrate ventrally. In contrast to fish and amphibians (Krotoski et al., 1988), migration of trunk NCC in amniotes is segmented and restricted to the anterior part of the somites (Gammill et al., 2006). There are also species-specific differences in the spatiotemporal specification of the melanoblast lineage. In avian and zebrafish dorso-lateral migration is delayed with respect to ventral migration (Krispin et al., 2010; Raible et al., 1992; Serbedzija et al., 1989). The molecular mechanism that controls this spatiotemporal lineage segregation is not fully understood, although some interesting studies suggest an important role for the NC specifier FoxD3 in regulating these decisions (Nitzan et al., 2013). In the mouse, trunk NCC migrate along both pathways at the level of the cervical region as early as e8.5, but melanoblasts pause in a region between the NT and the somites called the Migration Staging Area (MSA) (Weston, 1991) until e10.5, and only begin to colonize the epidermis at e11.5-12 (Serbedzija et al., 1990; Wilson et al., 2004).

0.5.4 Putative vertebrate NC gene regulatory network (NC-GRN)

NC development has been extensively studied in zebrafish, chick and frog models, but much less in the mouse and human species (Betancur et al., 2010; Betters et al., 2010; Stuhlmiller and Garcia-Castro, 2012; Thomas et al., 2008). The curiosity about the understanding of the origin and evolution of the NC has also prompted the analysis of NC development in the jawless vertebrate lamprey (Nikitina and Bronner-

Fraser, 2009; Sauka-Spengler et al., 2007). In all these models, NC forms in accordance with a stepwise mechanism involving 1) specification of the neural plate border, 2) specification of the NC, 3) emigration from the NT and colonization of the embryo, 4) homing and differentiation (for reviews see (Betancur et al., 2010; Green and Bronner, 2013; Meulemans and Bronner-Fraser, 2004; Sauka-Spengler and Bronner-Fraser, 2008; Stuhlmiller and Garcia-Castro, 2012)). The information from the study of the regulatory interactions that control each of these steps for the cranial NC population has been assembled in a putative NC-gene regulatory network (GRN) (Betancur et al., 2010; Stuhlmiller and Garcia-Castro, 2012).

As represented in Figure 0.11A, the NC-GRN starts with the induction of NCC at the neural plate border by the integration of signaling inputs from posteriorizing (canonical Wnt, FGF), Delta/Notch and intermediate levels of BMP coming from surrounding tissues such as the underlying mesoderm as well as the adjacent neural and non-neural ectoderm (Betancur et al., 2010; Stuhlmiller and Garcia-Castro, 2012). Combination of low/intermediate levels of BMP and canonical Wnt signals is essential for NC induction (Bang et al., 1999; Garcia-Castro et al., 2002; Liem et al., 1995). Integration of the signaling input induces gene expression of a regulatory kernel composed essentially of the transcription factors Pax3/7, Msx1/2, and Zic1/Zic2 which specify the neural plate border. Studies in *Xenopus* suggest that combination of Pax3 and Zic1 is sufficient to trigger the NC developmental program (Milet et al., 2013; Sato et al., 2005), whereas other studies in *Xenopus* and lamprey place *Msx1* as one of the earliest genes induced and acting upstream of other neural plate border specifiers such as *Pax3* (Monsoro-Burq et al., 2005; Nikitina et al., 2008). Subsequent cooperation between the neural plate border specifiers and signaling input activate gene expression of a second group of *bona fide* NC specifier transcription factors such as Id, cMyc, FoxD3, the SoxE members Sox9/10, Ets-1 and other genes (Bronner, 2014). NC specifiers control the survival and behavior of the NC population, including the next steps of EMT/delamination from the dorsal NT,

migration via stereotyped pathways and differentiation (Stuhlmiller and Garcia-Castro, 2012).

Comparative analysis of the expression and function of key players of the NC-GRN in several vertebrate models (including the lamprey) as well as basal chordates such as amphioxus and *Ciona* suggest a high conservation of the signaling input and the neural plate border specifier module across chordates (Bronner and LeDouarin, 2012; Green and Bronner, 2013; Holland, 2009; Shoguchi et al., 2008). The NC specifier module is highly conserved in vertebrates and lamprey, but is absent at the neural plate border in basal chordates (except for the expression of *Snail*). Since amphioxus lacks a bona fide NC, it is believed that co-option of homologs of vertebrate NC specifiers into the neural plate border was an important step in evolution of the NC (Bronner and LeDouarin, 2012; Green and Bronner, 2013).

On the other hand, it is well recognized that the proposed vertebrate NC-GRN is rudimentary and lacks important players (Bronner and LeDouarin, 2012). Indeed, in spite of the high conservation of the signaling inputs and neural plate border specifier module, it is still not well understood whether regulatory interactions are direct or not, how signaling inputs are integrated into cis-regulatory modules of *Pax*, *Msx* and *Zic* genes. Knowledge about the existence of feedback loops, cross-regulation within the network, is also rudimentary. Importantly, recent studies suggest that post-transcriptional and epigenetic modifications are also involved in the spatiotemporal control of NC formation (Bronner, 2014). The relative ease of performing gene perturbation analyses in *Xenopus*, zebrafish and chick has allowed a lot of progress during the last years in the understanding of how the NC-GRN is wired, but very little is known about this in mammals.

Taken together, all this evidence stress the idea that, although many elements of the NC-GRN are conserved across chordates (Green and Bronner, 2013), differences are

expected in the way the signaling inputs are transmitted and the network is integrated at the transcriptional level, between species and along the A-P axis.

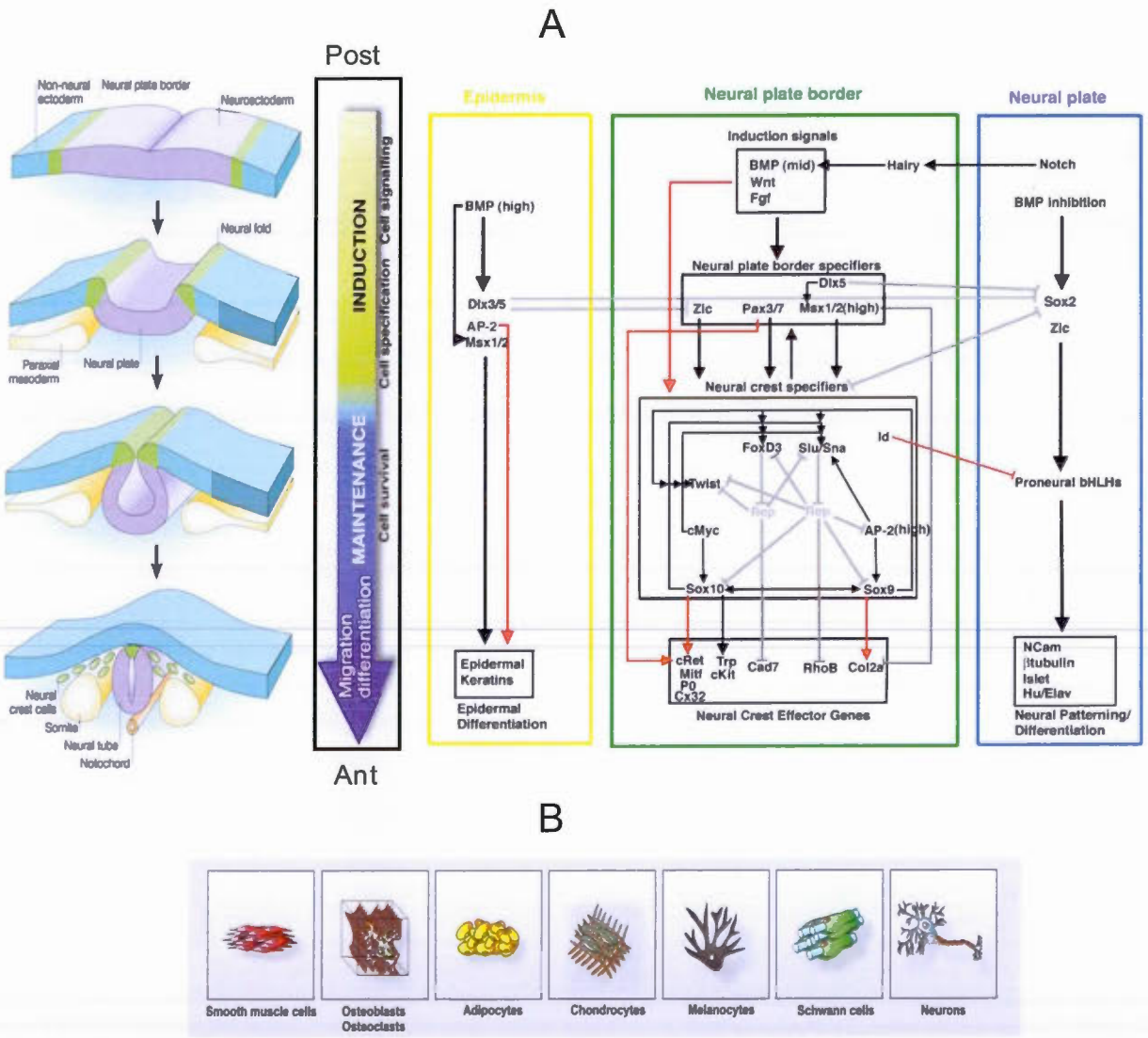


Figure 0.11 Schematic representation of neural crest (NC) development, the putative vertebrate NC gene regulatory network and NC derivatives.

(A) Representation of the different steps of NC development at the tissue level (left panel, taken from (Gammill and Bronner-Fraser, 2003)) in parallel to the proposed putative vertebrate gene regulatory network that controls these steps (right panel, taken from Meulemans and Bronner-Fraser, 2004). NCC are induced at the neural plate border (denoted

in green, left panel) upon the integration of signaling inputs in this region that turn on the expression of a first set of genes called neural plate border specifiers. In a second phase the induction signals cooperate with the neural plate border specifier module to maintain the NC progenitors and turn on the expression of definitive NC specifiers. These last control NC survival and behavior by regulating expression of NC effector genes. Once the neural plate roles into a neural tube, the NC (in green) localise to the dorsal neural tube, delaminate from this region to colonize the embryo and differentiate into several cell types (represented in B, adapted from Kaltschmidt lab web page).

0.5.5 Neural plate border specifiers Pax3/7, Msx1/2, Zics

0.5.5.1 *Pax3/7* genes

Pax3 and Pax7 belong to the Paired box (Pax) family of transcription factors (Goulding et al., 1991). The *Pax* family comprises nine members in mice and human (Pax1-9) (Wang et al., 2008a). *Pax* are important developmental genes generally involved in the control of cell lineage specification and maintenance of pluripotency of stem cell or progenitor cell populations. It has been shown that *Pax* genes fulfill these roles by controlling proliferation, inhibiting apoptosis or terminal differentiation of the cell populations in which they are expressed (Wang et al., 2008a). *Pax* genes normally exhibit restricted spatiotemporal expression patterns during embryogenesis, with expression normally turning off as tissues differentiate. Aberrant *Pax* expression in differentiated tissues is associated with development and progression of several types of cancer such as renal tumours, lymphoma, medullary thyroid carcinoma, rhabdomyosarcoma and melanoma (Wang et al., 2008a). It is proposed that *Pax* genes promote cancer development via the same mechanisms by which they control embryonic development: i.e: stimulation of cell proliferation, self-renewal and resistance to apoptosis (Lang et al., 2007).

It has been proposed that during evolution, *Pax3* and *Pax7* emerged by duplication of a common ancestral *Pax3/7* gene (Holland et al., 1999). Consistent with their common evolutionary origin these paralogs exhibit overlapping expression patterns in the developing nervous system and paraxial mesoderm during embryonic

development, as well as extensive structural and functional homology (Goulding et al., 1991). Both genes are partially co-expressed in the dorsal NT from the mesencephalon to the spinal cord regions as well as in the dermomyotome compartment of somites from early to midgestation (Relaix et al., 2004). These overlapping expression patterns correlate with similar important functions for both paralogs in the control of patterning and differentiation of the central nervous system, NC and paraxial mesoderm. Pax3 and Pax7 are implicated in the D-V patterning of the developing spinal cord. Indeed, both genes are expressed in commissural neurons where they redundantly prevent ventral neuronal identity (Mansouri and Gruss, 1998).

Of the two paralogs, Pax3 seems to play critical roles in early NT and NC development in the mouse. Homozygous *Pax3* loss-of-function severely affects dorsal NT closure, and leads to an almost complete absence of NC derivatives in the trunk region (Epstein et al., 1991; Lang et al., 2000). By contrast, *Pax7* mutation only generates late cephalic NC defects (Mansouri et al., 1996), suggesting that *Pax3* function is sufficient to initiate NT and NC development. Interestingly, gene replacement studies have clearly shown that *Pax7* can substitute for *Pax3* in these tissues (Relaix et al., 2004). This demonstrates functional overlap between the paralogs, which is also supported by the increase in severity of NT defects in *Pax3*/*Pax7* double homozygous mutants (Mansouri and Gruss, 1998).

The observation that endogenous *Pax7* expression cannot compensate *Pax3* loss-of-function is most likely due to differences in spatiotemporal expression patterns rather than function. Indeed, the overlap in gene expression is not perfect in the neuroectoderm along the A-P and D-V axis; *Pax7* expression is excluded from the most dorsal NT in the spinal cord and is absent in the posterior neural plate region, where *Pax3* expression is observed (Relaix et al., 2004). *Pax3* and *Pax7* are also differentially expressed in rhombomeres. *Pax3* is strongly expressed in rhombomeres

2 and 4, weakly in 1 and 3 and is absent from rhombomere 5, whereas *Pax7* exhibits the opposite distribution (Mansouri et al., 1996; Natoli et al., 1997). Expression of *Pax7* is up-regulated in the dorsal NT, somites and rhombomeres 2 and 4 in *Pax3* Splotch mutants, suggesting that Pax3 is involved in the restriction of *Pax7* expression (Borycki et al., 1999).

The Splotch mouse model and Pax3 neural function

The Splotch (Sp) mouse mutant was first reported by Russell (1947) as a spontaneous semi-dominant mutation causing white spotting of the belly, tail and feet in the heterozygous state, as well as spina bifida and early embryonic lethality (around e13.5) in homozygous mutants (Russell, 1947). Subsequently, other spontaneous (Sp^{d}) or radiation-inducible Splotch alleles (Sp^{r}, Sp^{1h}, Sp^{2h}) were identified that exhibited variable phenotypes, all mapping to chromosome 1 (see Epstein et al., 1991). The term Splotch comes from the characteristic white belly spot of heterozygous mutants. More detailed characterization was then provided by Auerbach (1954) and others, who have also reported the presence of several phenotypes revealing the pleitropic role of Pax3 in embryonic development. Defects include excencephaly, meningocele, curly tail, absence of limb muscles, severely reduced number or absence of NC derivatives (such as melanocytes, spinal and enteric ganglia), failure of cardiac outflow tract septation (the cause of embryonic lethality) as well as defects in thymus, thyroid, parathyroid development and a developmental delay leading to smaller size (Auerbach, 1954; Conway et al., 2000; Epstein et al., 1991; Lang et al., 2000).

Pax3 is located on chromosome 1 in mouse and maps to chromosome 2q35 in human (Goulding et al., 1991; Wang et al., 2008a). The murine *Pax3* cDNA sequence (1437 bp) was first isolated and characterized by Dr. Peter Gruss's lab in 1991. *Pax3* encodes a 56 KDa protein (479 amino acids) containing four structural domains that include two DNA binding domains: (a paired domain and a paired-type

homeodomain), an octapeptide motif and a transactivation domain (Goulding et al., 1991). The amino acid sequence of Pax3 is highly conserved between mouse and human, exhibiting 98% of sequence identity (Barber et al., 1999). Seven alternative splice forms of *Pax3* transcripts (*Pax3*a-e; *Pax3*g-h isoforms) have been reported in mouse and human and proposed to exhibit differential transactivating activity, biological function and expression patterns (reviewed in Wang et al., 2008a). The molecular basis of the Splotch mutation was revealed in 1991 by Douglas J. Epstein and co-workers from P. Gros lab (Epstein et al., 1991). By analyzing the genomic DNA of the Sp^{2h}/Sp^{2h} homozygous embryos the authors found a 32 nucleotide deletion within the paired domain region of *Pax3* which results in a truncated mRNA and protein with affected DNA binding function. Further characterization of the spontaneous Sp mutant showed the presence of mutation in the 3' AG splice acceptor site within intron 3 of the *Pax3* gene that results in abnormal splicing and generation of four aberrant *Pax3* mRNA coding for truncated Pax3 proteins (Epstein et al., 1993). Thus, the commonality of all Splotch mutations is the formation of aberrant transcripts that leads to a non functional Pax3 protein.

In humans, heterozygous *PAX3* mutation result in Waardenburg (WS) syndrome types I and III, a NC-related disorder (Tassabehji et al., 1992). Patients with WS I exhibit melanocyte-related abnormalities such as hair and skin hypo-pigmentation (a characteristic white forelock), hearing loss as well as craniofacial defects; WS III is characterized by musculo-skeletal abnormalities of the upper limb (Read and Newton, 1997). Children with heterozygous mutations in *PAX3* may also exhibit cardiac defects. As in the mouse, homozygous *PAX3* mutation is most likely lethal (Ayme and Philip, 1995). WS can also present associated with Hirschprung disease (or aganglionic megacolon), another NC-related disorder characterized by the absence of enteric ganglia in the colon region of the gut (Amiel et al., 2008). Although *PAX3* has not yet been identified as a Hirschprung-associated gene, *Pax3* Splotch homozygous mice exhibit severe aganglionosis and Pax3 has been shown to

interact with the Hirschprung-associated genes *Sox10* and *Ret* (described below). On the other hand, a recent study has revealed association of *PAX3* mutations with spina bifida in humans (Agopian et al., 2013). Altogether, these observations highlight that *Pax3* is an essential developmental gene in mouse and human. The *Pax3* Splotch model has been an invaluable tool to understand the molecular mechanisms by which Pax3 control NT and NC development in human (Figure 0.12).

Pax3 and the control of NT closure

Spina bifida is 100 % penetrant in Splotch homozygous mutants. Although several mechanisms have been proposed for the *Pax3*-dependent regulation of NT closure, how exactly the NT fails to close in Splotch homozygotes is not well understood (Copp et al., 2003; Zhao et al., 2014). It has been proposed that Pax3 controls apoptosis in the dorsal NT by stimulating degradation of the pro-apoptotic and tumor suppressor factor p53 (Wang et al., 2011). Indeed, excessive apoptosis and elevated p53 protein levels have been detected in the open NT of *Pax3* Splotch homozygous mutants, and inactivation of p53 can restore normal neurulation in these mutants (Pani et al., 2002; Phelan et al., 1997). It has been shown that Pax3 regulates p53 protein levels independently of transcription by physically interacting with p53 and the ubiquitin ligase Mdm2 via its DNA binding domains, whereas the non functional Pax3 Splotch protein fails to interact (Wang et al., 2011). All this work provides convincing mechanistic data of an anti-apoptotic role of Pax3 in NT closure. However, apoptosis appears to not be affected in ß-catenin conditional mutants that exhibit reduced *Pax3* expression levels and spina bifida (Zhao et al., 2014). Other studies propose that Pax3 may regulate cell proliferation and organization of the neuroepithelium. Indeed, it has been shown that the Splotch neuroepithelium exhibits reduced cell proliferation (Wilson, 1974), and is abnormally arranged with disorganized cells and increased intercellular spaces (Morris, 1983). These observations are consistent with other work suggesting that Pax3 regulates cell

surface properties in the dorsal NT in cell-autonomous manner as well as induces cell aggregation (Mansouri et al., 2001; Wiggan et al., 2002).

Pax3 and NC development

Pax3 is essential for NC development. Splotch homozygous mutants exhibit a severe reduction in the number of NCC emigrating from vagal and trunk regions and a complete loss of NCC emigrating from the caudal trunk (Serbedzija and McMahon, 1997). Anterior to the otic vesicle there are no major differences in the number, emigration and differentiation of the *Pax3*-defective NC population. The progressive reduction in the number of NCC from rostral to caudal levels in Splotch mutants results in a progressive decrease to complete lack of NC derivatives including cardiac NCC and melanoblasts as well as sympathetic, dorsal root and enteric ganglia (Auerbach, 1954). Importantly, these trunk NC defects are correlated with a lack of the NC specifier FoxD3, suggesting that Pax3 regulates trunk NC development at least in part via regulation of *FoxD3* gene expression (Dottori et al., 2001). Consistent with this, *FoxD3* expression is not impaired in areas of normal NC development (cranial levels) in Splotch homozygous embryos, most likely as a result of compensation by Pax7.

Pax3 is expressed in pre-migratory and early migrating NCC as well as in somites and several hypotheses have been raised regarding the mechanisms by which Pax3 controls NC development. At first, it was proposed that the NC defects in Splotch mutants were non-cell-autonomous and rather caused by the absence of *Pax3* expression in the somites, which are known to modulate the migratory behavior of NCC. This hypothesis was supported by heterograft experiments and the finding that mouse Splotch homozygous NCC when grafted into the chick embryo NT were able to migrate normally and follow appropriate migratory pathways (Serbedzija and McMahon, 1997). However, this non-cell-autonomous hypothesis has been refuted by the demonstration that transgenic re-expression of Pax3 in the dorsal NT and NC,

but not in somites, rescues the NC defects of Splotch homozygous embryos (Li et al., 1999).

Considerable work has supported a cell-autonomous and essential role for Pax3 in the regulation of the number of pre-migratory NC progenitors but not their migratory properties. Indeed, labeling of premigratory NCC and orthotopic transplantation studies have confirmed that homozygous Sp^{2h} mutant cardiac NCC are capable of migrating along normal pathways, although in considerably reduced number (Chan et al., 2004). More precisely, *Pax3* conditional knockout experiments and analysis of the cardiac population have revealed that Pax3 function is required very early during induction of cardiac NC progenitors but no longer once these cells emigrate from the NT (Olaopa et al., 2011). These and other studies then propose that the *in utero* lethal conotruncal heart defects of *Pax3* Splotch homozygous mutants is caused by a reduction in the number of cardiac NC progenitors and thus a deficient number of their derivatives in the outflow tract septum (Conway et al., 2000; Epstein et al., 2000; Olaopa et al., 2011).

The analysis of melanocyte development in Splotch homozygous mutants has equally revealed that Pax3 is required for the expansion of the pool of melanoblast precursors in the dorsal NT, but not for their migration. Indeed, transgenic tracing of the homozygous Splotch melanoblast populations has revealed a considerable reduction in the number of melanoblast in the mutant but these are equally capable to migrate along the dorsolateral pathway and arrive to their destination (Hornyak et al., 2001). Thus, the pigmentation defects of Splotch mutants are most likely due to a reduction in the pool of melanoblasts required to colonize the skin.

The number of enteric NC progenitors is considerably reduced in homozygous Splotch embryos and, consistent with this, Splotch enteric NCC fail to colonize the gut, leading to absence of enteric ganglia beyond the stomach (Lang et al., 2000). These defects can be rescued by re-expression of transgenic Pax3 in the dorsal NT

and NC thus demonstrating that Pax3 also acts cells autonomously in the regulation of enteric ganglia formation. Expression of the Ret tyrosine kinase receptor, which as described above is an essential regulator of enteric ganglia development in mouse and human, is considerably reduced in Sp/Sp mutants (Lang et al., 2000). It has been shown that Pax3 directly activates a *Ret* enhancer in cooperation with Sox10. *Ret* expression is detected in the developing enteric nervous system after enteric NCC have reached the gut. The spatiotemporal expression pattern of *Pax3* is thus consistent with a role in the initiation of *Ret* expression in enteric neural progenitors, but this regulation cannot explain the early reduction of NC-derived progenitors that enter the gut in Sp/Sp mutants.

In summary, as a neural plate border specifier Pax3 is essential for induction and expansion of the pre-migratory NC population. Loss of *Pax3* function as seen in the Sp/Sp mutant leads to a reduction of the number of pre-migratory NCC, the severity of this phenotype increases along the A-P axis and correlates with a reduction in gene expression of the NC specifier *FoxD3* (Dottori et al., 2001; Serbedzija and McMahon, 1997). Remaining NCC normally proliferate and are not impaired in their capacity of migration or differentiation (Conway et al., 2000; Hornyak et al., 2001; Olaopa et al., 2011; Serbedzija and McMahon, 1997). The precise molecular mechanism by which Pax3 controls expansion of the pre-migratory NC population is not well understood. Apoptosis is normally observed in the dorsal NT (Copp et al., 2003). It is likely that in this apoptotic environment Pax3 function is required cell-autonomously to control survival and/or proliferation of NC progenitors. The transcriptional activity of Pax3 is also essential for the specification of several NC-derived lineages. For example, Pax3 function is required to initiate gene expression of key regulators of either enteric neural progenitors (*Ret*) or melanoblasts (*Mitf*) (Bondurand et al., 2000; Lang et al., 2000), as well as to repress gene expression of the negative regulator of cardiac NC development *Msx2* (Kwang et al., 2002).

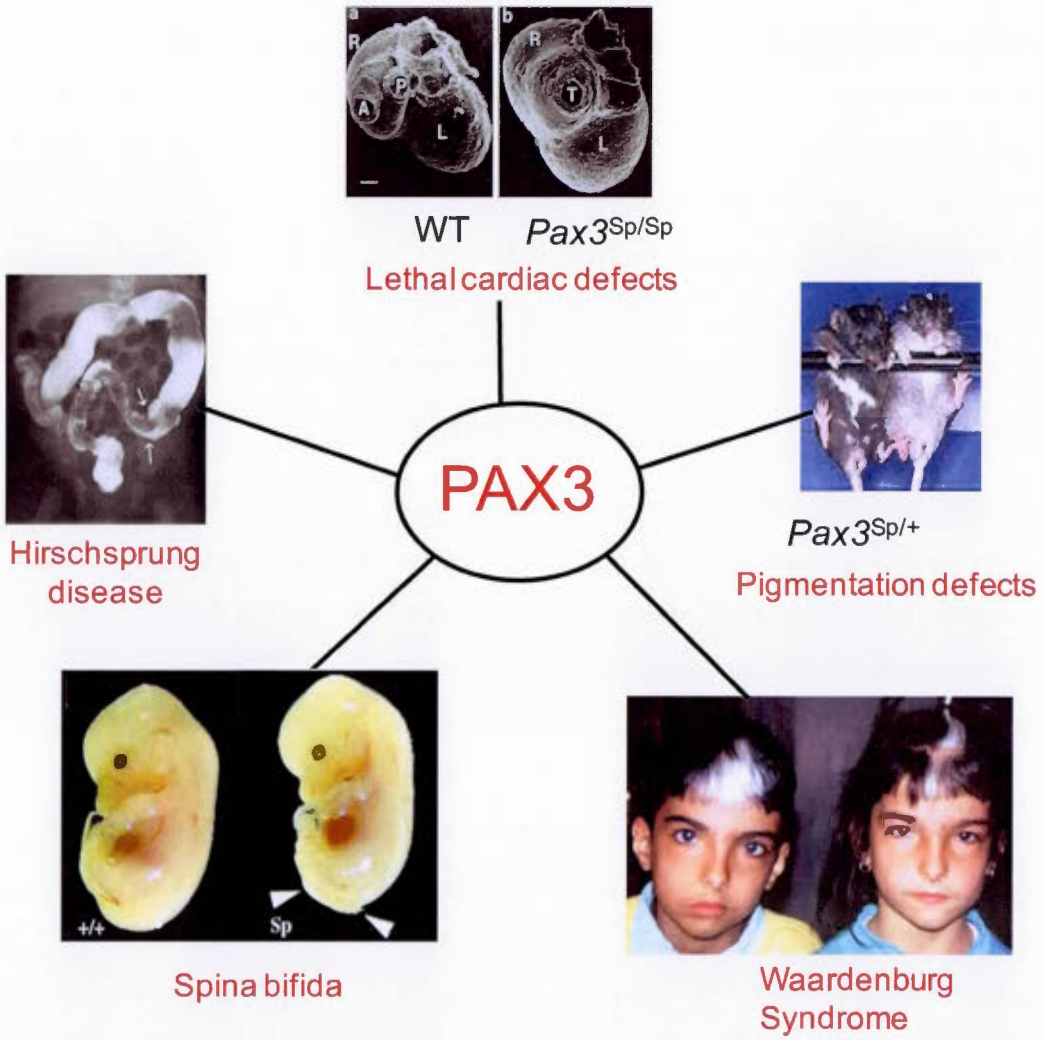


Figure 0.12 Neurocristopathies and neural tube defects associated with loss of Pax3 function in mouse and human.

Regulation of posterior Pax3 neural expression domain by signaling inputs

In the mouse, at the onset of expression (~e8.5), *Pax3* transcripts are detected at the lateral edges of the neural plate extending from the midbrain to sligthly more posteriorly of the hindbrain-spinal cord boundary in the open PNP (Goulding et al., 1991). This typical posteriorly and lateraly restricted pattern of *Pax3* expression is also observed in other vertebrate models including *Xenopus* and zebrafish (Monsoro-

Burq et al., 2005; Moore et al., 2013). After NT closure, the domain of *Pax3* expression is restricted to the dorsal NT where Pax3 marks pre-migratory NCC as well as other dorsal neural progenitors (commissural neurons and interneurons). Species-specific differences are also observed as seen for example in chick embryos where *Pax3* is initially broadly expressed within the whole PNP and then becomes progressively dorsally restricted in the closed NT (Bang et al., 1997; Goulding et al., 1993; Liem et al., 1995). The spatiotemporal pattern of *Pax3* expression is thus consistent with this gene being a target of A-P and D-V patterning signals.

Considerable work has focused on the understanding of the molecular mechanisms that establish and refine the spatiotemporal expression pattern of *Pax3* in the neuroectoderm. Pioneer transplantation assays performed in chick first established that signals from the notochord and floor plate are involved in the dorsal restriction of *Pax3* expression in the spinal cord (Goulding et al., 1993). Of particular relevance was the observation that removal of the notochord caused a ventral expansion of *Pax3* expression in the NT, whereas grafting a supernumerary notochord in the dorsal domain caused a considerable reduction in *Pax3* expression. Shh was then identified as the notochord-derived repressive signal based on the fact that the *Pax3* expression domain is shifted ventrally in *Shh*-null mouse mutants and that Shh represses *Pax3* expression in chick dorsal NT explants (Chiang et al., 1996; Liem et al., 1995). In parallel, studies from T.M. Jessell's group evidenced that BMP signaling coming from the epidermis is involved in the induction and dorsal maintenance of the *Pax3* expression domain (Liem et al., 1995). Particularly, this work showed that contact with epidermal ectoderm or incubation with BMP4 and BMP7 induce *Pax3* expression in explants derived from propective ventral and intermediary regions of the closing neural plate, and upregulate *Pax3* expression in explants derived from prospective dorsal regions of the closing neural plate. Taken together, all this work implicated counteracting dorsalizing BMP signaling coming from the epidermis and

ventralizing Shh signals coming from the notochord and floor plate in the dorsal restriction of *Pax3* expression in the spinal cord.

Subsequent studies were aimed at the identification of the signaling inputs that induce *Pax3* expression in the PNP. The answer was provided by Goulding's lab in a series of studies performed using ectodermal explants isolated from *Xenopus* (animal caps) and chick embryos (Bang et al., 1999; Bang et al., 1997). These studies showed that *Pax3* expression is induced in the PNP in response to posteriorizing signals (FGF, RA and Wnt) coming from the late GO and paraxial mesoderm, with canonical Wnt signaling playing an essential role in *Pax3* induction. Indeed, whereas treatment with bFGF or RA induces *Pax3* expression in neuralized *Xenopus* animal caps (treated with Noggin), *Xenopus*/chick sandwich explants assays showed that induction of *Pax3* expression by the paraxial mesoderm depends of Wnt signals but not FGF or RA, as it is blocked by a dominant negative (dn) form of Wnt8 but not by inhibition of FGF or RA signals (Bang et al., 1999; Bang et al., 1997). Thus, according to these studies FGF and RA signals can induce *Pax3* expression in neuralized animal caps but they are dispensable for this regulation in vivo. However, there must be species-specific differences since perturbation of FGF signaling does affect *Pax3* expression levels in the zebrafish PNP border (Garnett et al., 2012). Moreover, this study in the zebrafish model evidenced that FGF acts redundantly with canonical Wnt in the amplification of *Pax3* expression, with *Pax3* transcript levels in the PNP border decreasing more when both posteriorizing signals are knocked down than when each Wnt or FGF signaling is knocked down alone (Garnett et al., 2012). On the other hand, these studies have shown that intermediary levels of BMP signaling position the *Pax3* expression domain in the PNP border as overexpression of the BMP ligand *bmp2b* caused a medial shift in the *Pax3* expression domain in the trunk region (Garnett et al., 2012).

In summary, this work can be assembled in a model in which *Pax3* expression is initiated at the PNP by integration of posteriorizing canonical Wnt signaling coming from surrounding tissues (and probably FGF and RA acting redundantly) and intermediate levels of BMP signaling coming from the epidermis (Bang et al., 1999; Garnett et al., 2012; Liem et al., 1995). The distribution of Wnt and BMP signals in the mouse posterior embryo supports this model. Canonical Wnt signaling is broadly active in the caudal embryo but cannot account for *Pax3* lateral restriction (Maretto et al., 2003). By contrast, BMP signaling activity is restricted to the lateral edges of the open posterior neural folds (Tozer et al., 2013; Yang and Klingensmith, 2006; Ybot-Gonzalez et al., 2007a) and particular thresholds are required to induce *Pax3* expression (Garnett et al., 2012; Liem et al., 1995). More recent work in zebrafish and chick propose a role for Shh signaling in the restriction of *Pax3* expression to the lateral borders of the PNP (Moore et al., 2013). Later on, in the closed NT, *Pax3* dorsal restriction is maintained by a positive input from dorsalizing BMP signaling coming from the epidermis and a negative input from ventralizing Shh signaling coming from the notochord and floor plate that prevents *Pax3* expression in the ventral NT (Chiang et al., 1996; Goulding et al., 1993).

Neural Pax3 cis-regulatory modules (CRMs) and transcriptional regulators

Once known that posteriorizing Wnt, dorsalizing BMP and ventralizing Shh are the major signaling inputs that set up the A-P and D-V pattern of *Pax3* expression in the caudal neuroectoderm, the next obvious question is how positional information is integrated at the cis-regulatory level? What are the molecular effectors of these signals? This is yet not fully understood in the case of *Pax3*. A good way to solve this question is to identify the minimal regulatory sequence in the genome able to recapitulate the expression pattern of the gene and then scan this sequence for transcription factors binding sites, the last players in the signaling molecular cascade. This approach has led to the identification of two redundant regulatory regions at the *Pax3* locus: one located in the proximal promoter and the other in the fourth intron

(Degenhardt et al., 2010; Natoli et al., 1997). Each of these regions is enriched in evolutionarily-conserved CRMs and contain sufficient positional information to recapitulate the spatiotemporal induction and dorsal restriction of murine *Pax3* expression in the neuroectoderm (Degenhardt et al., 2010; Milewski et al., 2004; Natoli et al., 1997). A considerable number of *Pax3* regulators with different spatiotemporal expression patterns that bind and regulate these sequences have been identified, thus demonstrating that *Pax3* regulation is complex and dynamic along the A-P axis.

The 1.6 kb proximal promoter regulatory region was identified by Natoli et al., in 1997 in a screen for *Pax3* regulatory sequences *in vitro* using RA-differentiated murine embryonal carcinoma P19 cells and was confirmed using mouse transgenic reporter assays (Natoli et al., 1997). This region directs specific *Pax3* expression in the dorsal NT of e9.5 embryos from the hindbrain down to the tailbud and recapitulates the expression pattern of *Pax3* in rhombomeres. However, broad ectopic *Pax3* expression is observed in the whole tailbud, suggesting that the 1.6 kb promoter lacks the regulatory elements that normally prevent *Pax3* expression in this zone. Although the identity of a *Pax3* repressor in the tailbud has not yet been revealed, regulatory elements are likely to be located upstream of the 1.6 kb region and within the 14 kb proximal promoter region, since no ectopic reporter expression is observed in the tailbud of transgenic embryos bearing the 14 kb construct (Natoli et al., 1997). Regulatory elements that direct *Pax3* neural expression anterior to the hindbrain and in the somites are also located in this region (Milewski et al., 2004; Natoli et al., 1997).

The proximal 1.6 kb *Pax3* promoter not only recapitulates *Pax3* neural expression patterns in pre-migratory NCC and dorsal NT but also yields enough functional *Pax3* expression levels. Indeed, this regulatory region has been used for re-expressing *Pax3* in homozygous Splotch mutants and effectively rescues NT closure and NC

development including normal cardiac outflow tract septation and enteric ganglia formation in Splotch mutants (Lang et al., 2000; Li et al., 1999). The 1.6 kb promoter has been used to engineer a Cre recombinase line called P3ProCre, (Li et al., 2000) widely employed for conditional gene inactivation and lineage tracing of dorsal NT and NC derivatives and interestingly also metanephric mesenchyme as well as a subset of skeletal muscle lineage derivatives (Chang et al., 2004; Jarad and Miner, 2009; Li et al., 2000).

Deletion analysis of the 1.6 kb *Pax3* proximal promoter narrowed down the minimal NC enhancer (NCE) to a block of 647 bp containing two CRMs named NCE1 and NCE2, each of approximately 250 bp in length (Milewski et al., 2004). Sequence comparisons have revealed that NCE1 and NCE2 are highly conserved between mouse and human. This study also proposed that both elements are required to control *Pax3* expression in the hindbrain, most dorsal NT and trunk NC, as deletion of either NCE1 or NCE2 completely abrogate reporter expression in transgenic embryos. However, interpretation of these results is limited by the small number of transgenic embryos analysed. Indeed, whereas deletion of NCE2 abolished reporter expression in n=3 transgenic embryos, the consequence of deletion of NCE1 was only tested in n=1 embryo (Milewski et al., 2004).

After the identification of NCE1 and NCE2 as the minimal *Pax3* CRMs, was the finding of DNA binding factors that regulate these elements. It has been shown that both NCE1 and NCE2 bear functional binding sites for the neural-specific Pou class III family of transcription factors Brn1 and Brn2, which are necessary for the activity of the 1.6 kb *Pax3* proximal promoter along the rostro-caudal axis in e9.5 transgenic embryos (Pruitt et al., 2004). It has also been shown that NCE1 contains functional Pbx regulatory elements that bind Pbx1/Meis and Pbx1/Hox (paralogs group 1-5) complexes. Transgenic and perturbation analyses have shown that Pbx1 is critical for the activity of the 1.6 kb *Pax3* proximal promoter in rhombomeres and pre-migratory

cardiac NCC (Chang et al., 2008; Pruitt et al., 2004). On the other hand, NCE2 contains a functional binding site for the transcription factor Tead2 which is critical for the activity of the whole NCE in transgenic mice (Milewski et al., 2004). Intriguingly, although overexpression of a dominant negative Tead2 fusion protein in the dorsal NT and NC affects *Pax3* expression in the dorsal NT and dorsal root ganglia as well as development of these structures (Milewski et al., 2004), inactivation of Tead2 does not affect *Pax3* expression (Kaneko et al., 2007). Moreover, three consensus Lef/Tcf binding sites have been found in the 1.6 kb *Pax3* proximal promoter: two of them are located in the NCE1 region and the other in the minimal promoter region. The later has been shown to be required for the activity of the 1.2 kb *Pax3* proximal promoter in tissue culture and is occupied by ß-catenin (Zhao et al., 2014).

Surprisingly, targeted deletion of the whole NCE in mice using homologous recombination failed to affect *Pax3* expression (Degenhardt et al., 2010). This led to the identification of two other evolutionarily conserved regions (ECR2 and ECR4), located in the 4th intron, that act redundantly with the NCE (Degenhardt et al., 2010). Studies in zebrafish and chick embryos have revealed that both enhancers – also called IR1 and IR2 (Garnett et al., 2012) or CNE3 and CNE1 (Moore et al., 2013), respectively – integrate different signaling cues and act coordinately to control induction and dorsal restriction of *Pax3* expression in the NT. Based on studies in the zebrafish, it has been proposed that CNE3 (ECR2 or IR1) mediates induction of *Pax3* expression in the caudal neuroectoderm, as it directly integrates an inductive input from canonical Wnt signaling and a repressive input from the Shh target Nkx6.1 (Moore et al., 2013). CNE3 also responds to BMP and FGF perturbations in the zebrafish but it is still unclear how these signals are integrated at the cis-regulatory level (Garnett et al., 2012). On the other hand, CNE1 does not respond to canonical Wnt, but can equally respond to FGF and BMP perturbations in zebrafish (Garnett et al., 2012). Cis-regulatory analysis performed in zebrafish and chick embryos further

suggest that CNE1 is implicated in the maintenance of *Pax3* dorsal restriction and supports a direct auto-regulatory loop from Pax3/7 as well as a direct positive input from neural specific SoxB members (Moore et al., 2013). Sox3 and Sox11 transcription factors have been proposed as candidates in this regulation according to ChIP-Seq data from stem cell derived neuronal progenitors showing occupancy of these transcription factors on CNE1 (Bergsland et al., 2011).

0.5.5.2 *Msx* genes

Msx1 together with its paralogs *Msx2* and *Msx3*, compose the three vertebrate members of the *Msx/Msh* gene family of homeobox-containing transcription factors. This family is highly conserved and found from basal animals species such as the sea anemone to more complex species like humans (Finnerty et al., 2009; Takahashi et al., 2008). Mouse *Msx* members were identified by Holland in 1991 as homologous to the *Drosophila melanogaster* muscle segment homeobox (*Msh*) gene (Holland, 1991). The 60 amino acid long helix-turn-helix homeodomain as well as an Engrailed Homology 1 (Eh1) motif located at the N-terminus, which is involved in transcriptional repression, are highly conserved between Msx1, Msx2 and Msx3 proteins (Takahashi et al., 2008). Outside these two domains there is low sequence similarity between Msx proteins (Catron et al., 1996; Takahashi et al., 2008). Msx transcription factors have a well recognized role as repressors (Alappat et al., 2003; Catron et al., 1996), and appear to act independently of their DNA binding properties to directly interfere with the transcriptional machinery (Mehra-Chaudhary et al., 2001; Shetty et al., 1999) or by forming a repressor complex with Groucho proteins via the Eh1 motif (Rave-Harel et al., 2005; Takahashi et al., 2008). Growing evidence also suggests that Msx1 and Msx2 can act as transcriptional activators and can fulfill this role in both DNA-independent and –dependent manner by interacting with cofactors or via binding to a common homeodomain binding site, as is the case of Atoh1 which has been recently reported to be a positive direct target of Msx1 and Msx2 in the dorsal spinal cord (Duval et al., 2014).

Of the mammalian *Msx* genes, *Msx1* and *Msx2* are the best characterized in part because of their importance in craniofacial morphogenesis. During early mouse embryonic development, both genes are co-expressed in the dorsal NT and NC along the AP axis, where *Pax3* also is expressed, as well as in the craniofacial primordia, pharyngeal arches and limb buds (Catron et al., 1996). However, the spatial domain of *Msx1* in these tissues is broader, compared to that of *Msx2*. Other differences in expression, for example the unique expression of *Msx1* in the otic vesicle as well as the wider *Msx1* expression in the distal regions of the nasal processes and the maxillary and mandibular components of the first branchial arch, may account for the differences in the phenotypes of Msx1 and Msx2 loss-of-function mutants (described below). *Msx3* is exclusively expressed in the dorsal NT with an anterior limit in the rostral rhombencephalon (Wang et al., 1996). *Msx3* expression transiently overlaps with *Msx1* and *Msx2* at E9.5 but then becomes excluded from the roof plate where Msx1 and Msx2 are localized (Wang et al., 1996). The consequences of inactivating *Msx3* have not yet been reported.

Both *Msx1*- and *Msx2*-null mutant mice exhibit craniofacial abnormalities; however the phenotypes are not the same. *Msx1* mutants have tooth defects, cleft palate and die shortly after birth (Satokata and Maas, 1994), whereas *Msx2* mutants are viable and exhibit defects in development of the skin, teeth, jaws and skull vault (Satokata et al., 2000). Functional overlap between Msx1 and Msx2 is clearly evidenced by the increase in severity of craniofacial phenotypes, emergence of new phenotypes as well as embryonic lethality at late gestation when both genes are inactivated (Ishii et al., 2005). Compound *Msx1/2* mutants exhibit severe defects in cranial and cardiac NC development, including mis-patterning and reduced cranial ganglia size, dysmorphogenesis of pharyngeal arch derivatives, malformation of cranial bones and cardiac outflow tract defects. *Msx1/2* double mutants also exhibit NT defects such as exencephaly and spina bifida (the latter at low penetrance), severe defects in neuronal patterning as well as in limb development (Duval et al., 2014; Ishii et al., 2005;

Lallemand et al., 2005). Since no trunk NC-related phenotypes have been described so far in single or compound *Msx* mutants, it has been proposed that *Msx* genes are not critical for trunk NC development (Simoes-Costa et al., 2012). However, a conclusion about the role of Msx in mouse NC development must await for the results of concomitant inactivation of *Msx1*, *Msx2* and *Msx3*, as functional overlap of *Msx3* and embryonic lethality of *Msx1/2* double compound mutants may have precluded the observation of other NC phenotypes.

Consistent with the phenotypes of single mouse mutants, loss-of function of *MSX1* in humans is associated with selective tooth agenesis, cleft palate, cleft lip, and nail dysplasia (Jumlongras et al., 2001; van den Boogaard et al., 2000; Vastardis et al., 1996). Mutations in *MSX2* cause Familial Parietal Foramina characterized by persistent unossified areas in the skull vault, as well as a defect known as craniosynostosis characterized by premature fusion of the calvarial bones at the sutures (Wilkie et al., 2000; Wuyts et al., 2000).

Msx1 and NC development

In the mouse, *Msx1* is expressed at the neural plate border region and then in pre-migratory and early migrating NCC (Catron et al., 1996; Houzelstein et al., 1997). However, the complete understanding of the roles of Msx1 in mouse NC development has been hampered by functional overlap between Msx members (Catron et al., 1996; Ishii et al., 2005). Although single and compound *Msx1/Msx2* mouse mutants do not display overt defects in early NC formation (Ishii et al., 2005), studies in *Xenopus* and chick embryos have clearly shown that Msx1 is involved in NC specification. In *Xenopus* embryos, forced expression of *Msx1* is sufficient to induce NCC as determinated by the induction of the NC markers *Snail2*, *Slug* and *FoxD3*, whereas inhibition of Msx1 function by using a dominant negative form of Msx1 represses NC induction (Monsoro-Burq et al., 2005; Tribulo et al., 2003). Morpholino-mediated inhibition of Msx1 function leads to pigmentation defects

(Monsoro-Burq et al., 2005). On the other hand, overexpression of *Msx1* by electroporation in the chick NT induces the expression of the NC marker *Dlx* (Liu et al., 2004). Epistasis and cis-regulatory analyses have placed *Msx1* downstream of BMP, FGF and canonical Wnt and upstream of *Pax3*, *FoxD3* and *Snail2* in the NC-GRN (Bang et al., 1999; Bang et al., 1997; Betancur et al., 2010; Liem et al., 1995; Monsoro-Burq et al., 2005; Nikitina et al., 2008; Simoes-Costa et al., 2012; Tribulo et al., 2003). Msx1 is involved in the control of cell proliferation, differentiation and survival in several tissues. Characterization of the NC phenotypes in *Msx1/2* double compound mouse embryos suggests a delay in migration of the cranial NC population (although it is unclear whether it is caused by reduced production, delamination or migration), patterning defects and excessive apoptosis in NC-derived cells that contribute to the cranial ganglia and first pharyngeal arch (Ishii et al., 2005).

Regulation of Msx1 expression and CRMs

In vitro, *ex vivo* and *in vivo* analysis in several vertebrate species suggests that *Msx1* expression is regulated by positional cues from dorsalizing BMP, ventralizing Shh and posteriorizing canonical Wnt and FGF signaling pathways. How these inputs are coordinated at the cis-regulatory level to establish the complex spatiotemporal expression pattern of *Msx1* is still unclear. *Msx1* spatiotemporal expression pattern strikingly overlaps with that of *Pax3* in the caudal neuroectoderm. Both genes are early markers of the PNP border and spinal cord D-V patterning. Consistent with the similar expression and functions, *Pax3* and *Msx1* are regulated by similar mechanisms. *Msx1* is a well known target of BMP signaling (Alvarez Martinez et al., 2002; Suzuki et al., 1997). Studies in *Xenopus* and chick embryos suggest that, like *Pax3*, the induction of *Msx1* expression in the PNP border requires intermediate levels of BMP signaling coming from the epidermal ectoderm as well as canonical Wnt signaling coming from the paraxial mesoderm (Bang et al., 1999; Liem et al., 1995; Tribulo et al., 2003). *Msx1* neural induction is also dependent on FGF8 signaling which has been proposed to act in parallel to canonical Wnt in this process

(Monsoro-Burq et al., 2005). In the early spinal cord, the dorsal domain of *Msx1* expression is maintained by integration of dorsalizing BMP signaling coming from the epidermis and ventralizing Shh signaling from the notochord and floor plate (Liem et al., 1995).

The analysis of the regulatory regions responsible for *Msx1* expression can be traced back to twenty years ago when Kusuoka et al. analyzed the sequence of the 1280 bp *Msx1* proximal promoter and reported the presence of a considerable number of potential cis-regulatory elements (Kuzuoka et al., 1994). These include putative binding sites for heterodimers of RA and retinoid X receptors RAR:RXR, six Tcf1 binding sites, two Msx1 binding sites, as well as several binding sites for ubiquitous transcription factors including three Ebox and three SP1 motifs (Kuzuoka et al., 1994). Other work has also identified a consensus Smad binding site near the transcription start site (TSS), which appears to mediate activation by BMP signaling as it is directly bound by Smad proteins and is required for transcription of *Msx1* in tissue culture (Alvarez Martinez et al., 2002). The proximal promoter also contains three Nuclear factor-kappa Beta (NF-kB) binding sites that mediate transcriptional activation by FGF signaling and are required for activation of *Msx1* expression in the chick limbs (Bushdid et al., 2001).

Co-transfection and deletion analysis of the 5 kb region upstream of the transcription start site (TSS) revealed that the 1282 bp proximal promoter is sufficient and necessary to activate *Msx1* transcription in the mouse myoblast cell line C2C12 (Takahashi et al., 1997). Subsequent deletion of the 1282 bp region identified a minimal promoter, located -165/+106 bp relative to the TSS. The *Msx1* minimal promoter lacks a TATA binding sequence but contains one functional binding site for specificity protein 1 (Sp1) transcription factor which is essential for its activation in co-transfection analysis (Takahashi et al., 1997). Indeed, it has been proposed that activation of *Msx1* transcription requires Sp1 binding to its GC sequence motif,

which tethers and stabilizes a multisubunit complex comprising TATA box binding protein (TBP), Sp1 and the co-activator CBP/p300 on the minimal promoter; (Shetty et al., 1999). Msx1 is known to repress its owns expression. However, although there is a functional *Msx1* consensus binding site in the minimal promoter, Msx1 apparently act independently of DNA binding via physically interacting and squelching the transcriptional complex (Shetty et al., 1999).

Mackenzie et al., performed transgenic reporter analysis in mice and found that the regulatory elements responsible for the complex and dynamic spatiotemporal expression pattern of *Msx1* are concentrated 4.9 kb upstream of the transcriptional start site (TSS) (MacKenzie et al., 1997). Deletion analysis of this region further narrowed down the Msx1 regulatory regions to two CRM, one named distal enhancer (DE, 240 bp long, located -4670 to -4420 bp upstream of the TSS), and the other called proximal enhancer (PE, 78 bp long, located -2630 to -2553 bp upstream of the TSS). According to the analysis of a LacZ reporter in e9.5-12.5 transgenic embryos, DE drives expression in developing nasal processes and the first and second pharyngeal arches, whereas PE controls expression in a wide range of tissues including the third pharyngeal arch, roof plate, trunk somitic dermamyotome, limb bud mesenchyme, optic vesicle, second arch, genital ridge and epiphysis (MacKenzie et al., 1997). PE bears AP-2, ETS, a functional Smad binding site as well as a consensus TCF binding site which is directly bound by TCF4 *in vivo* in response to canonical Wnt signaling, and is required for the activity of PE in e10.5 embryos (Alvarez Martinez et al., 2002; Miller et al., 2007).

Although Mackenzie's study failed to show enhancer activity for the *Msx1* proximal promoter (MacKenzie et al., 1997), a parallel study revealed that the basal promoter and particularly sequences -165/+106 bp from the TSS drive transgenic reporter expression in the craniofacial tissues particularly in developing tooth mesenchyme of e10.5 transgenic embryos (Takahashi et al., 1997). Information responsible for *Msx1*

expression in the cardiac NC population is located in an evolutionarily conserved cardiac NC-specific enhancer (KE). Located 40 kb 5' of the *Msx1* TSS, KE is regulated by cAMP mediated PKA signaling (Miller et al., 2008). A study of the regulatory region that controls the spatiotemporal expression of *Msx1* in the trunk NC has not been reported.

0.5.5.3 *Zic* genes

Originally isolated from the murine cerebellum (Aruga et al., 1994), the zinc finger of the cerebellum (Zic) vertebrate gene family encodes multifunctional transcription factors that play critical roles in several processes during embryogenesis; (for reviews see Houtmeyers et al., 2013; Merzdorf, 2007). Just to mention a few, Zic proteins have been involved in the regulation of neurogenesis, skeletal patterning, myogenesis, left-right asymmetry; cerebellar patterning and retinal neuron development. It is believed that Zic proteins control specification and proliferation of the tissues in which they are expressed (Merzdorf, 2007).

Zic genes are the vertebrate homologues of the *Drosophila melanogaster* pair rule gene *odd-paired* (opa) (Aruga et al., 1996). In most vertebrate species the Zic family comprises five members (Zic1-5), except for zebrafish that contains seven *Zic* genes. Zic/opa proteins are all characterized by the presence of five tandem Cys2His2-type zinc fingers that exhibits high sequence homology to the zinc finger domain of Gli/Ci proteins (Houtmeyers et al., 2013) (Figure 0.13). The ZF1 is different from the others, in the sense that the two Cys residues are separated by a long amino acid sequence. The ZF domain is highly conserved between Zic members, and is essential for Zic function. Indeed, this domain has been involved in DNA binding, protein-protein interaction and also in nuclear shuttling of the protein (Houtmeyers et al., 2013; Merzdorf, 2007).

It is not yet well understood which ZF precisely mediates DNA binding or protein-protein interaction, although structural and mutational analyses suggest that ZF 3-5 are important for DNA recognition (Koyabu et al., 2001). These canonical ZF can also mediate protein binding, but in that case it appears that protein-protein interaction blocks DNA recognition. Indeed, it has been shown that Zic counteract Gli-mediated transcriptional activation and thus Shh signaling and *vice versa* via direct physical interaction between ZF 3-5 of both proteins (Koyabu et al., 2001). On the other hand, structural analysis suggest that Zic ZF 1-2 are not canonical DNA binding fingers and may be involved in protein-protein interaction (Houtmeyers et al., 2013).

Zic proteins bear other regions of high sequence homology, apart from the ZF domain. These include a ZF N-terminally conserved domain (ZF-NC), which as the name indicates is located immediately N-terminal of the ZF domain. The function of ZF-NC is currently unknown (Aruga et al., 2006). Zic1, Zic2 and Zic3 and Opa present an N-terminally located domain, called Zic/Odd-paired conserved motif (ZOC) (Aruga et al., 1996) that is involved in protein-protein interactions (Mizugishi et al., 2004). ZOC is also part of the transcription regulatory domain, mapped to the N-terminal region, which can act as either activator or repressor depending of the cellular context (Mizugishi et al., 2004). On the basis of the presence of the ZOC motif, Zic proteins are divided into two structural subclasses: the class A includes Zic1, Zic2 and Zic3. This subgroup not only presents the ZOC motif but also a conserved ZF 1. The class B (Zic4 and Zic5), lacks the ZOC box and exhibit differences in ZF 1 structure (Houtmeyers et al., 2013). All Zic proteins are also characterized by the presence of stretches of conserved alanine and histidine residues located outside the ZF domain (Houtmeyers et al., 2013). Mutations in these low complexity regions are associated with human diseases and result in decreased DNA binding activity (Brown et al., 2001).

Because of the high conservation of the ZF domain, co-expressed Zic members exhibit functional overlap and can potentially bind and recognize similar target genes (Merzdorf, 2007). *In vitro* DNA binding assays suggest that Zic preferentially bind to Gli-binding sequences (5'GGGTGGTC3'), consistent with the notable homology in the ZF domain of both families (Mizugishi et al., 2001). However, the Zic-binding sequence found in the regulatory elements of the few Zic target genes reported so far are largely different (Aruga, 2004; Merzdorf, 2007). The commonality of Zic target sequences is the presence of three to five Gs or Cs in the core of the sequence. Based on this, it is proposed that Zic bind GC-rich target sequences rather than to a consensus element (Aruga, 2004). Zic can act both as direct activators or repressors of gene expression (Ebert et al., 2003; Salero et al., 2001; Yang et al., 2000), depending on the cellular context and interaction with cofactors (Mizugishi et al., 2001). Zic can also regulate gene expression independently of DNA binding and by acting as cofactors. The best well characterized binding partners of Zic are Gli proteins. It has also been shown that Zic ZF domain directly binds to TCF proteins and inhibits activation of canonical Wnt-TCF target genes (Pourebrahim et al., 2011).

Zic genes exhibit a particular genomic arrangement (Aruga et al., 2006; Houtmeyers et al., 2013; Merzdorf, 2007) (Figure 0.13). In most vertebrate genomes the five *Zic* genes are arranged in two tandem pairs composed of *Zic1-Zic4* and *Zic2-Zic5* while *Zic3* is unpaired. Each pair is composed by members of different subgroups, share the same regulatory region and consequently, exhibit similar expression patterns and function. Interestingly, in teleost genomes such as those of zebrafish and *Fugu* there are seven *Zic* genes, arranged in three pairs *zic1-zic4*; *zic2a-zic5*; *zic3-zic6*; with *zic2b* being unpaired. Phylogenetic analysis performed by Aruga's group proposed that this particular genomic arrangement arose by *cis* and *trans* duplications of a common ancestor, resulting in eight zic genes (4 pairs on different chromosomes). During evolution, most vertebrate genomes like mice, humans and *Xenopus* lost one pair and the partner of *Zic3* (Aruga et al., 2006).

Expansion in copy number during evolution is proposed to account for functional overlap, differences in expression patterns, as well as specific functions of Zic members. The spatiotemporal expression pattern of *Zic* genes in the mouse neuroectoderm only partially overlaps across the A-P and D-V axes (reviewed in (Merzdorf, 2007)). Zic2, Zic3 and Zic5 are the only members expressed in the neuroectoderm at early somite stage (e8.5), when NC induction begins. At e9.5, *Zic2* and *Zic5* are co-expressed in the dorsal NT along the entire A-P axis and in the open PNP; *Zic1* expression pattern largely overlaps with that of *Zic2* and *Zic5* but is absent from the most caudal region. *Zic3* expression is restricted to the forebrain region, whereas *Zic4* is expressed in the dorsal midline of the brain and anterior trunk NT. The domain of expression across the D-V axis of the spinal cord also varies between Zic members (see (Merzdorf, 2007)). *Zic1*, *Zic2* and *Zic5* dorsal domains overlap in the most dorsal NT, although the expression domain of *Zic1* expands more ventrally relative to that of *Zic2* and *Zic5*; *Zic4* is excluded from the roof plate; *Zic3* is very weakly, or not, expressed in the spinal cord (Nagai et al., 1997). Of note, *Zic* genes are also differentially expressed in mesoderm-derived tissues during embryogenesis, including somites (Merzdorf, 2007).

It is well recognized that signaling inputs from BMP, canonical Wnt, FGF and Shh pathways are involved in the establishment and refinement of *Zic* neural expression domain (Aruga, 2004; Garnett et al., 2012; Merzdorf, 2007). However, as extensively stated in this work, it is not well understood how these A-P and D-V positional cues are integrated at the cis-regulatory level. Several studies in *Xenopus* and zebrafish embryos suggest that canonical Wnt and FGF act redundantly in the induction of *Zic* neural expression (Garnett et al., 2012; Monsoro-Burq et al., 2003). Particularly, cis-regulatory analyses in the zebrafish have shown that the *Zic2a-Zic5* pair and *Zic3* are directly downstream of canonical Wnt signaling (Garnett et al., 2012; Nyholm et al., 2007). BMP signaling has been implicated in the refinement of the mediolateral expression domain of the *Zic* genes and maintenance of dorsal restriction in the

closed NT. Indeed, it has been shown that BMP has a dual role in regulation of *Zic* neural expression pattern. Studies in *Xenopus* and zebrafish have well demonstrated that induction of *Zic* neural expression requires attenuation of BMP signaling, with BMP antagonists Noggin, Chordin, and Follistatin playing a key role in this process (Aruga, 2004; Garnett et al., 2012). *Xenopus* animal cap assays suggest that Msx1 may be an intermediary in the repression of *Zic* genes by high levels of BMP signaling and act to exclude *Zic* expression in the lateral ectoderm (reviewed in (Aruga, 2004; Merzdorf, 2007)).

The dorsal restriction of *Zic* expression in the NT is established by counteracting gradients of BMP and Shh signaling. It has been shown that perturbation of BMP signaling by misexpression of BMP4/7 in the chick dorsal NT positively regulates *Zic* expression (Aruga et al., 2002b). Work in several species including the mouse has demonstrated that *Zic* genes are negative targets of ventralizing Shh signaling. Indeed, *Zic1* and *Zic2* expression domain expands ventrally in the NT of mutant mice with notochord degeneration or in *Shh*-null mutant mice (Brown et al., 2003; Nagai et al., 1997). It is expected that repression of *Zic* genes by Shh morphogen in the ventral spinal cord is indirect, as Shh signaling transduction culminates in the stabilization of the activator form of the ZF domain Gli transcription factors (Dessaud et al., 2008). However, the Zic CRMs and transacting factors that mediate the Shh repressive input are unknown.

Loss-of-function mutations of every mouse *Zic* genes has been reported and lead to different phenotypes, suggesting both functional overlap and functional specificity between members. Deletion of *Zic1* results in cerebellar malformations and axial skeletal defects (Aruga et al., 1998; Aruga et al., 1999). Mouse null mutants for *Zic4* do not exhibit any overt phenotype, but in combination with *Zic1* null mutation leads to more severe cerebellar phenotypes suggesting that Zic1 and Zic4 redundantly control cerebellum development (Blank et al., 2011). Consistent with the

observations in mice, heterozygous deletion of the *ZIC1-ZIC4* pair in human causes a congenital cerebellar malformation known as Dandy Walker syndrome (Grinberg et al., 2004). Loss-of-function of *Zic3* in mice (Bent tail null mutant) and human leads to defects in left-right asymmetry including heart malformations as well as NTDs such as exencephaly (Carrel et al., 2000).

Reduced expression or complete loss-of-function of mouse *Zic2*, as seen in embryos homozygous for the hypomorphic $Zic2^{Kd}$ allele ($Zic2^{Kd/Kd}$), or the Kumba allele ($Zic2^{ku/ku}$), respectively, results in NTDs such as spina bifida and exencephaly, as well as a defect in forebrain development known as holoprosencephaly (Elms et al., 2003; Nagai et al., 2000). Heterozygous mutation of human *ZIC2* causes holoprosencephaly as well (Brown et al., 1998). The mouse Kumba allele is an ENU-induced missense mutation that affects the second Cys residue of the 4^{th} ZF and prevents DNA binding (Elms et al., 2003; Nolan et al., 2000). Homozygous *Zic2* Kumba embryos also exhibit defects in hindbrain patterning and a severe decrease in the number of NCC along the whole A-P axis (Elms et al., 2003). Characterization of the NC phenotypes of Kumba mutants suggests delayed NC induction and severe decrease in *FoxD3* expression. Homozygous Kumba NCC normally proliferate and migrate (Elms et al., 2003). Thus, the Kumba mouse model is reminiscent of *Pax3* Splotch mutants. Indeed, similar to *Pax3* Splotch heterozygous mice, heterozygous Kumba mice also exhibit pigmentary anomalies such as the characteristic white belly spot as well as occasional spina bifida and curly/kinked tail (Elms et al., 2003).

Consistent with the phenotypes of *Zic2* mutants, loss-of-function of the paired *Zic5* gene also causes NTDs, particularly exencephaly, and severely affects the number of NCC leading to craniofacial defects (Inoue et al., 2004). However, *Zic5* mutation only affects induction of the cranial NC population, suggesting that *Zic2* compensates for *Zic5* loss at trunk levels. The phenotypes of *Zic2/Zic5* double compound mutants have not been reported so far.

In summary, of the five Zic members only Zic2, Zic3 and Zic5 are involved in the control of NT closure. The penetrance of NTDs is low, between 10-20% (Aruga, 2004). Zic2 appears to have unique roles in the control of spinal cord neurulation. In mice, the *Zic2-Zic5* pair has been involved in the early control of NC development, particularly in NC induction. Mutation in *Zic2* or *Zic5* similarly leads to a severe decrease in the number of NCC, but not in their migratory properties. However, these genes appear to contribute differentially to NC development along the A-P axis. *Zic2* is essential for NC induction along the entire A-P axis, whereas *Zic5* is critical for formation of the cranial NC population. Although a similar role for *Zic1*, *Zic3* and *Zic4* has not been reported in mice, all five *Zic* members have been involved in the regulation of *Xenopus* NC development (reviewed in (Aruga, 2004; Merzdorf, 2007)).

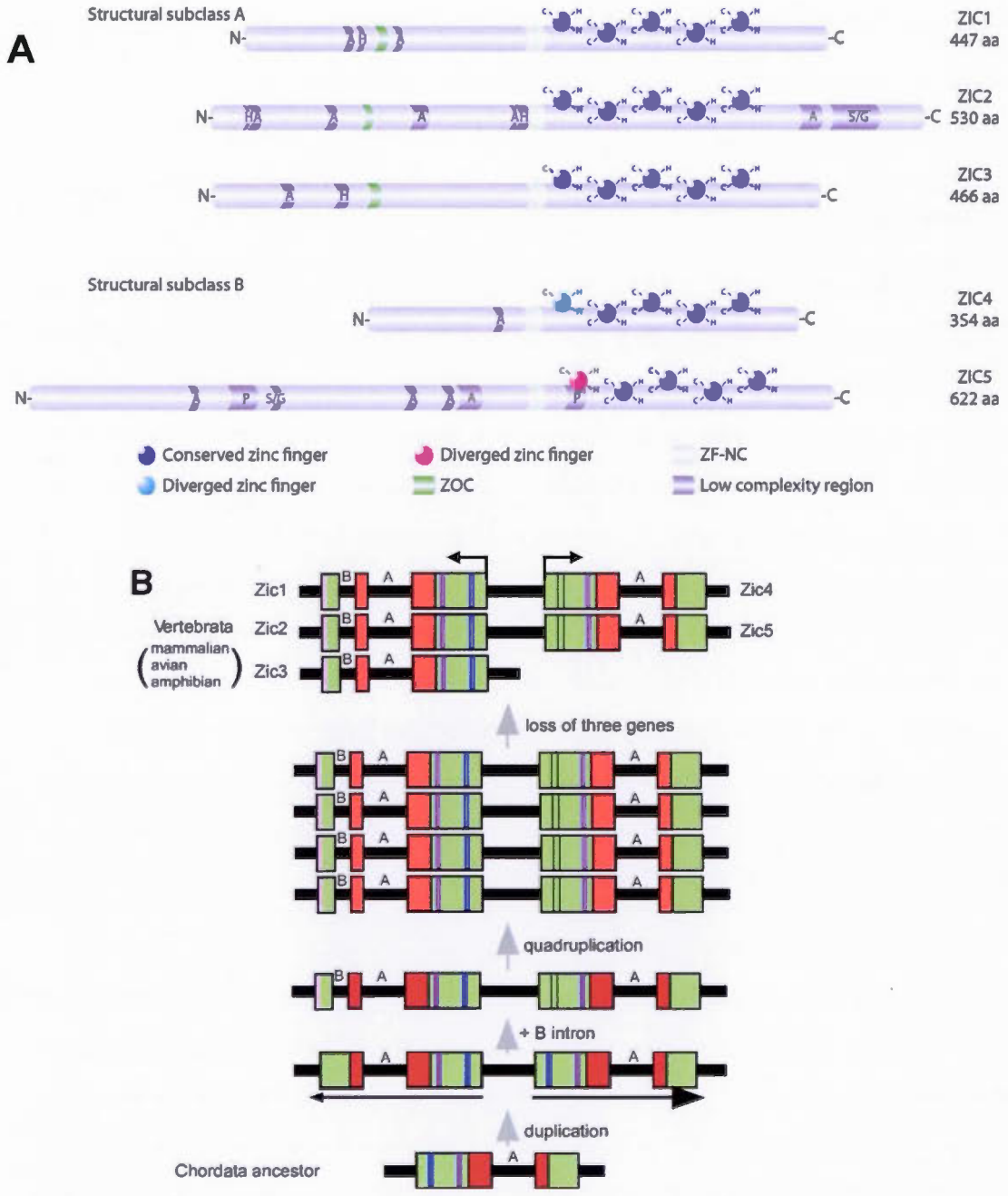


Figure 0.13 Zic protein structure and genomic arrangement of *Zic* genes in several vertebrate species.

(A) ZIC1, ZIC2 and ZIC3 proteins contain a ZOC motif (depicted in green) and belong to the subclass A, whereas ZIC4 and ZIC5 lack this motif and belong to the subclass B. Source of

the figure: Houtmeyers et al., 2013. Note that the first zinc finger is also highly conserved in members of the subclass A. (B) Proposed model that explains the establishment and genomic arrangement of vertebrate *Zic1-5* genes. Taken from Aruga et al., 2006. Protein coding regions are indicated by boxes, black horizontal bars indicate intronic and regulatory regions. Green: protein coding regions, red: zinc finger domain, dark blue: ZOC motif; purple: ZF-NC.

0.5.6 FoxD3

Forkhead Box D3 (*FoxD3*, also known as *Hfh2* or *Genesis*) encodes a transcription factor belonging to the winged helix or forkhead gene family (Labosky and Kaestner, 1998; Sutton et al., 1996). Isolated from a murine embryonal carcinoma cDNA library (Sutton et al., 1996), *FoxD3* encodes a 465 amino acid protein containing a conserved helix-turn-helix DNA-binding domain known as the winged helix or forkhead box (100 amino acid in size), as well as several motifs commonly involved in transcriptional regulation. FoxD3 lacks the common regulatory motifs II, III and IV of most members of the winged helix family and rather shares sequence similarity with a particular subfamily of winged helix genes that are commonly expressed in embryonic cells (Sutton et al., 1996). FoxD3 has long been recognized as a transcriptional repressor (Sutton et al., 1996), however increasing evidence suggests that it can act either as a context-dependent repressor or activator of gene expression (Abel et al., 2013; Pan et al., 2006).

FoxD3 is a stemness factor that was originally named *Genesis* because it was found to be expressed only in mouse ES cells or their malignant equivalents (including P19 cells) (Sutton et al., 1996). *FOXD3* is also a marker of human ES cell lines (Hanna et al., 2002). Further analysis in the mouse embryo revealed that *FoxD3* is expressed during early embryogenesis in multipotent cell populations particularly in the epiblast, and then in the NC (Labosky and Kaestner, 1998). *FoxD3* is a *bona fide* NC marker. In several vertebrate species such as mouse, *Xenopus*, chick and zebrafish, *FoxD3* is one of the first markers of premigratory NCC. *FoxD3* is also expressed in multiple migrating NC-derived lineages except melanoblasts. Consistent with the

observations in embryonal carcinoma cells, *FoxD3* is also down-regulated during NCC differentiation.

In the mouse embryo, *FoxD3* transcripts are broadly distributed in the ExE and epiblast at e6.5-7.5, and then become restricted to pre-migratory NCC by e8.0-e8.5 (Labosky and Kaestner, 1998). *FoxD3* and *Pax3* expression patterns extensively overlap. At e8.5, *FoxD3* transcripts are detected in the lateral borders of the head folds, in the recently closed dorsal NT at the level of the hindbrain-spinal cord boundary. Like *Pax3*, expression of *FoxD3* is excluded from rhombomeres 3 and 5. Induction of *FoxD3* then follows the temporal course of trunk NC induction and is observed in premigratory NCC of the recently closed dorsal NT at e9.5. At this stage, *FoxD3* expression is also observed in cranial sensory ganglia, the lens placode, and ventrally migrating trunk NCC. By e10.5, *FoxD3* is also expressed in postmitotic interneurons of the spinal cord. At e11.5, *FoxD3* expression is detected in a number of NCC derivatives, but extinguishes as cells differentiate.

Consistent with the particular expression of *FoxD3* in multipotent cells, loss-of-function studies in the mouse have demonstrated that *FoxD3* is an important stem cell factor required for the maintenance of pluripotency and self-renewal capacity of progenitor cell populations, including ICM cells and NCC (Hanna et al., 2002; Teng et al., 2008). *FoxD3* KO ES cell lines cannot be established. Mouse embryos die at early gastrulation (~e6.5) with a complete loss of epiblast cells and expansion of proximal ExE and endoderm (Hanna et al., 2002). Characterization of the mutants has shown that *FoxD3* is required for survival, proliferation and maintenance of both embryonic and trophoblast progenitors (Hanna et al., 2002; Tompers et al., 2005).

Conditional inactivation of *FoxD3* in the NC by using a Wnt1-Cre line results in a dramatic loss of NC progenitors and NC derivatives along the entire A-P axis (Teng et al., 2008). NC development is more severely affected at trunk levels, with mutant embryos exhibiting a complete lack of vagal and sacral enteric NC progenitors.

Consistent with these severe defects, *FoxD3* cKO mutants die perinatally with craniofacial malformations, including cleft face and palate, reduced peripheral nervous system and lack of the enteric nervous system. It was shown that, as in the case of epiblast cells, *FoxD3* is required early in NC development to control the survival and expansion of the pool of NC progenitors. In the absence of *FoxD3,* mouse NC progenitors appear to be correctly specified but most premigratory and early migrating NCC undergo apoptosis. The few NCC that survive, however, are able to migrate and differentiate normally. Mutant embryos also exhibit reduction in number of cardiac NC progenitors but effectively septate the outflow tract (Teng et al., 2008). More recently it was shown that *Pax3* and *FoxD3* genetically interact in the control of survival and maintenance of NC progenitors including cardiac NCC, with introduction of one Splotch allele to the *FoxD3* cKO background causing a complete loss in NC progenitors caudal to the first pharyngeal arch and increasing the penetrance of cardiac outflow tract defects to 100% (Nelms et al., 2011). Conditional *FoxD3*-null mutants also exhibit defects in melanocyte development (described below) (Nitzan et al., 2013).

Studies in other vertebrate species have corroborated the important role of *FoxD3* in maintenance of pluripotency of NCC. In chick, forced expression of *FoxD3* in migrating NCC inhibits differentiation and retains cells in a multipotent state (Dottori et al., 2001). Loss of *FoxD3* function in zebrafish leads to a considerable loss of NC derivatives due to failure in survival of NC progenitors (Lister et al., 2006; Montero-Balaguer et al., 2006). A role for *FoxD3* in promoting NC specification and delamination from the NT has also been shown in species different from mice (Cheung et al., 2005; Dottori et al., 2001; Kos et al., 2001; Sasai et al., 2001). In chick embryos, misexpression of *FoxD3* in the NT induces expression of the migratory NC markers HNK1 and Cad7, and increases emigration from the NT (Dottori et al., 2001). In *Xenopus* embryos, ectopic expression of *FoxD3* also increases expression of NC markers whereas expression of a dominant negative

FoxD3 affects NC specification and leads to loss of NC derivatives (Sasai et al., 2001).

FOXD3 is a marker of human embryonic stem cells (Arduini and Brivanlou, 2012). Adult human tissues do not normally express this stem cell factor. Consequently, upregulation of *FOXD3* expression has been associated with either promotion or inhibition of development of several types of cancers including neurofibromas, chronic lymphocytic leukemia and melanomas (Weiss et al., 2014b).

0.5.6.1 Regulators of *FoxD3* expression in the NC and CRMs

Gain- and loss-of-function studies, as well as epistasis analyses in *Xenopus*, lamprey and mouse embryos have placed *FoxD3* downstream of Msx1/2, Pax3/7, and Zic1 inputs in the NC-GRN (Dottori et al., 2001; Monsoro-Burq et al., 2005; Sato et al., 2005; Sauka-Spengler et al., 2007; Simoes-Costa et al., 2012; Tribulo et al., 2003). In *Xenopus* and lamprey embryos, knockdown of *Msx1*, *Pax3* or *Zic1* results in a decrease of *FoxD3* expression at the neural plate border, whereas misexpression of these genes induces ectopic *FoxD3* expression in *Xenopus* (Monsoro-Burq et al., 2005; Sato et al., 2005; Sauka-Spengler et al., 2007; Tribulo et al., 2003). It has been shown that induction of *FoxD3* by Pax3 and Zic1 requires further cooperation with canonical Wnt signaling in *Xenopus* (Sato et al., 2005).

In chick embryos, two evolutionarily conserved NC-specific enhancers (called NC1 and NC2) have been recently found that act complementarily to control the spatiotemporal expression pattern of *FoxD3* in the NC along the A-P axis (Simoes-Costa et al., 2012). NC1 is located approximately 20 kb upstream from the TSS, and is a cranial NC-specific enhancer. NC1 acts early to initiate *FoxD3* expression in the cranial NC region, with activity observed in premigratory and early migrating cranial NCC from the midbrain down to rhombomere 2 (R2). Deletion analysis narrowed down the minimal regulatory core region to an 80 bp fragment and mutational,

chromatin immunoprecipitation (ChIP) and perturbation analysis suggest that NC1 is directly regulated by Pax7, Msx1/2 and the cranial NC marker Ets1. On the other hand, NC2 is located approximately 44 kb upstream of the TSS, and is a trunk specific-enhancer that recapitulates initial *FoxD3* expression in premigratory and early migrating vagal and trunk NCC with an anterior limit of activity at the level of R6. NC2 is also responsible for *FoxD3* expression in R4 and acts later to control expression in a particular subpopulation of migrating cranial NCC. NC2 activity was also observed in migrating melanoblasts, suggesting that regulatory regions responsible for repression of *Foxd3* in the melanoblast population are located outside NC2. A 90 bp core region was identified as sufficient for the activity of NC2. Similar to NC1, NC2 is directly regulated by Pax7 and Msx1/2, however its trunk-specific activity is mediated by an additional direct input from Zic1. In summary, this work in chick embryos provided the first cues of how *FoxD3* NC expression is regulated at the *cis*-regulatory level, and suggests that region-specific mechanisms act to differentially regulate *FoxD3* expression in the cranial and trunk NC populations (Simoes-Costa et al., 2012). Whereas Pax3/Pax7 and Msx1/2 act via regulation of both NC1 and NC2, specific additional inputs from Ets1 on NC1 and Zic1 on NC2 appear to be critical to control the spatiotemporal expression of *FoxD3* in the NC.

Consistent with studies in lower vertebrates, *FoxD3* is also downstream of Pax3 in mice (Dottori et al., 2001). Phenotypes of *FoxD3* cKO mutant embryos are reminiscent of those of *Pax3* Splotch homozygous mutants mostly at trunk levels where Pax7 is not co-expressed, and *FoxD3* expression is severely decreased in caudal regions lacking trunk NC derivatives (Dottori et al., 2001). The phenocopy of cranial NC defects of *Msx1/2* double mutants also suggests that *FoxD3* is downstream of *Msx1/2* in mice (Ishii et al., 2005; Teng et al., 2008). How Pax3/7 and Msx1/2 inputs are integrated at the cis-regulatory level in the mouse is currently unknown.

A very recent study provided the first *cis*-regulatory analysis in mice and reported that the A-P patterning protein Hoxb5 regulates *FoxD3* expression via direct binding to its proximal promoter region (Kam et al., 2014a). Hoxb5 regulation is important for *FoxD3* expression and function in the trunk NC, as perturbation of Hoxb5 function in the NC by using a conditional Cre-Loxp-dependent dominant negative approach severely reduced the levels of *FoxD3* transcripts in the trunk region, and phenocopied apoptosis and considerable reduction in the number of migratory trunk NCC of conditional *FoxD3* mutants (Kam et al., 2014a; Kam et al., 2014b).

0.5.7 Melanocyte development

Melanocytes are a specialized and highly migratory NC-derived cell type that produces the pigment melanin (Rawles 1947, 1948). These amazing cells are not only responsible for coloring and photo-protecting the epidermis, hair and iris but also play important physiologic functions in development and homeostasis of the inner ear, meninges of the brain and heart (Lang et al., 2013; Levin et al., 2009; Pavan and Raible, 2012; Tachibana, 1999).

Melanocyte development involves several steps and begins with the specification of melanocyte precursors (melanoblasts) from NCC (Pavan and Raible, 2012). Melanoblasts are specified at all axial levels with the trunk NC population having a big contribution to this lineage (Bronner and LeDouarin, 2012). According to lineage tracing analysis in chick embryos, melanoblasts specification begins before emigration of NCC from the dorsal NT (Krispin et al., 2010). Molecular characterization of this population in the chick suggests that prospective melanoblasts downregulate expression of *FoxD3*, *Sox9* and *Snail2,* otherwise these factors inhibit melanogenesis and emigration (Nitzan et al., 2013).

Upon delamination from the NT, melanoblasts follow a stereotypic path beneath the ectoderm known as the dorso-lateral pathway, which is reserved for the melanocyte

lineage, so that almost every NC cell taking this pathway is fated to become a melanocyte (Pavan and Raible, 2012). Although the dorso-lateral pathway is the major migratory way for melanocytes that populate the skin and hair follicles, skin melanocytes can also arise from Schwann cell precursors that migrate along the ventral pathway (Adameyko et al., 2009). Other melanocyte populations such as those found in the inner ear and heart also arise from these Schwann cell precursors (Pavan and Raible, 2012).

Lineage tracing analysis in mice have shown that melanoblasts taking the dorsolateral pathway leave the dorsal NT as early as e8.5 (Serbedzija et al., 1990) and move to another resting site known as the Migration Staging Area (MSA) (Weston, 1991). Melanoblasts pause in this particular zone located between the NT and the somites where they are exposed to signals that direct expression of melanocyte-specific genes and dorso-lateral migration. In the MSA, specification into *bona fide* melanoblast begins with the expression of the early marker microphtalmia-associated transcription factor (Mitf) (Opdecamp et al., 1997). Soon after, melanoblasts also express other early markers including the melanogenic enzyme Dopachrome tautomerase (Dct) also known as tyrosinase-related protein-2 (TRP-2), as well as the receptor tyrosine kinase cKit (Hari et al., 2012; Opdecamp et al., 1997; Steel et al., 1992). Melanoblasts also expand in numbers in the MSA and leave this resting zone from e10.5 onwards to take the dorsolateral pathway and colonize the epidermis at e11.5-12 (Pavan and Raible, 2012 and references therein).

In subsequent steps, skin melanoblasts undergo extensive migration throughout the embryo, proliferate and further differentiate into mature melanocytes which also express the melanogenic enzymes Tyrosinase (Tyr) and Tyrosine-related protein-1 (Trp-1), develop special organelles known as melanosomes and acquire a dendritic morphology (Pavan and Raible, 2012; Steel et al., 1992). Melanosomes are used to synthesize the two forms of melanin: the brown/black eumelanin and red/yellow

pheomelanin whereas dendritic processes mediate the transfer of melanin to skin keratinocytes and hair follicles (Pavan and Raible, 2012). Some melanoblasts remain quiescent and form a population of melanocyte stem cells (MSCs) found within the bulge region of the hair follicle (Lang et al., 2013). This population shares several common features with melanoma stem cells and is characterized by the expression of *Pax3* and *Dct* (Osawa et al., 2005); mature melanocytes downregulate *Pax3*. MSCs provide the source for newly differentiating melanocytes during adulthood. Canonical Wnt signaling plays a key role in the control of proliferation and differentiation of melanocyte stem cells within the niche (see Lang et al., 2013).

0.5.7.1 Pax3, FoxD3, Sox9 and the melanocyte GRN

Melanocyte development is controlled by a gene regulatory network with Mitf being the earliest melanoblast marker and essential for specification of the melanocyte lineage (Opdecamp et al., 1997). Pax3 acts as a nodal point in the control of melanocyte development (Lang et al., 2005). Pax3 initiates the melanogenic gene regulatory network by directly activating *Mitf* expression in nascent melanoblasts in cooperation with Sox10 (Bondurand et al., 2000), whereas at the same time maintains melanoblasts in an undifferentiated state by forming a repressor complex with the Groucho protein Grg4 and Lef1 that prevents Mitf binding and activation of the *Dct* promoter (Kubic et al., 2008; Lang et al., 2005). *Dct* repression by Pax3 is relieved by canonical Wnt signaling: nuclear ß-catenin displaces the Pax3-repressor complex from the *Dct* enhancer and forms an activator complex together with Lef1 and Mitf (Lang et al., 2005). The Mitf-activator complex may function alone or in cooperation with Sox10 to transactivate *Dct* expression (Kubic et al., 2008). Mitf also regulates expression of important regulators of melanocyte development and survival, including the expression of genes encoding enzymes of the melanin biosynthetic pathway, such as *Tyr* and *Tyrp1* (Levy et al., 2006; Vance and Goding, 2004). Consistent with its role as a master regulator of melanocyte development, null mutations in *Mitf* in several species results in a complete loss of melanocytes (Pavan

and Raible, 2012). In humans, mutations in *MITF* are associated with Waardenburg syndrome Type 2, characterized by congenital hearing impairment and pigmentary anomalies in eye, hair and skin (Tassabehji et al., 1994). Similarly, deregulation in important components of the melanocyte GRN, such as Pax3, Sox10, Wnt signaling, Sox9, or the tyrosinase enzymes Dct, Tyr1, Tyrp1 also affect melanocyte development and lead to pigmentary anomalies (Guyonneau et al., 2004; Hornyak et al., 2001; Lane and Liu, 1984; Pavan and Raible, 2012; Qin et al., 2004).

Melanocyte development is also controlled by the interaction with environmental growth factors. Two important cell extrinsic factors for melanocytes are the Mast cell growth factor (*Mgf*) and Endothelin3 (*Edn3*). Mutation in these genes encoding these proteins, as seen in the Steel (Mgf^{Sl}) or lethal spotting ($Edn3^{ls}$) mice lead to defects in melanocyte development and pigmentary anomalies ranging from white spotting to total lack of coat pigmentation (Baynash et al., 1994; Copeland et al., 1990). Accordingly, expression in melanoblasts of the receptors for these signaling molecules, as is the case of cKit (the receptor for Mgf) and Endothelin receptor type b (Ednrb, the receptor for Edn3) is essential for their response to these environmental cues and normal survival, proliferation, guidance and differentiation (reviewed in (Pavan and Raible, 2012)). Consequently, mutations in *cKit* and *Ednrb* as seen in dominant spotting (Kit^{W}) and piebald (*Ednrb*) mice, also lead to pigmentary anomalies and other defects, including aganglionic megacolon in the case of the piebald mutants (Geissler et al., 1988; Hosoda et al., 1994).

FoxD3 is an essential regulator of melanocyte development (Curran et al., 2009; Dottori et al., 2001; Kos et al., 2001; Nitzan et al., 2013; Thomas and Erickson, 2009). Studies in chick embryos suggest that FoxD3 represses melanogenesis by two independent mechanisms: 1) it prevents melanogenic specification by inhibiting the binding of Pax3 to the *Mitf* promoter (Thomas and Erickson, 2009) and 2) it prevents dorso-lateral migration by repressing expression of the guidance receptor Ednrb2

(Nitzan et al., 2013). In chick embryos, misexpression of Ednrb2 in the dorsal NT is sufficient to switch the migration of neural progenitors into the dorsolateral pathway, but does not induce the melanogenic fate. Although it is yet unclear whether the FoxD3-Ednrb2 circuit is active in the mouse, the same authors showed that conditional *FoxD3* loss-of-function in the mouse NC using the Wnt1Cre line results in ectopic activation of *Mitf* in the dorsal NT and premature entrance of *Mitf*-expressing melanoblasts in the dorsolateral pathway at e10.5 (Nitzan et al., 2013). In summary, FoxD3 is involved in the regulation of the timing of melanocyte specification and entrance into the dorsolateral pathway.

Sox9 is expressed in premigratory trunk NCC, plays a key role in trunk NC survival and epithelial-mesenchymal transition and is downregulated after trunk NCC emigrate from the dorsal NT (Cheung et al., 2005). The study by Nitzan et al. in the chick model also revealed that Sox9 and Snail2 act upstream of FoxD3 and compose a gene regulatory network in the dorsal NT that controls segregation of neural and melanocyte lineages (Nitzan et al., 2013). *Sox9* expression is downregulated in melanoblast progenitors (Krispin et al., 2010) and it was shown that, like FoxD3, Sox9 represses melanoblast specification and pathfinding (Nitzan et al., 2013). When electroporated in the chick NT at late stage (35 somite stage), Sox9 prevented the expression of the melanoblast marker MC/1 and upregulated the glial marker myelin protein zero (MP0) but not the neural marker neurofilament in late emigrating NC-progenitors. Moreover, these late progenitors ectopically expressing Sox9 failed to express the guidance receptor Endrb2 and migrated ventrally instead of taking the dorsolateral pathway (Nitzan et al., 2013). However, other studies have shown that upregulation of *Sox9* in the dorsal NT at early stages (10 somite stage) promotes melanocyte differentiation including premature entrance into the dorsolateral pathway, probably due to its role in promoting NC specification at early stages (Cheung and Briscoe, 2003). Intriguingly, Odd Sex (Ods) transgenic mice which

overexpress *Sox9* in NCC and maintain high levels in melanoblasts from e11.5 onwards, exhibit hypopigmentation including a white belly spot (Qin et al., 2004).

0.6 Caudal-related homeobox (Cdx) genes. Expression pattern, regulation and function

0.6.1 The *Caudal* gene family

The *Caudal* gene family (*Cad* in arthropods/*Cdx* from Caudal-related homeobox in chordates) encodes homeodomain transcription factors belonging to the ancestral *ParaHox* gene cluster (Brooke et al., 1998; Pollard and Holland, 2000). *Caudal* gene products have essential and evolutionarily conserved roles in development of the posterior embryo in Bilaterians. In arthropods and vertebrates, loss-of-function of the *Caudal* homologues *Cad/Cdx* results in similar anterior-posterior (A-P) patterning defects, and severe posterior truncation (Copf et al., 2004; Faas and Isaacs, 2009; Macdonald and Struhl, 1986; Subramanian et al., 1995; van Rooijen et al., 2012). Consistent with their conserved roles, Caudal members present a highly conserved DNA binding homeodomain, and exhibit similar posteriorly restricted expression patterns in the developing embryo (Gamer and Wright, 1993; Macdonald and Struhl, 1986; Meyer and Gruss, 1993; Mlodzik and Gehring, 1987; Scott et al., 1989).

The first Caudal member (cad) was identified in *Drosophila* as an homeobox-containing gene by Mlodzik (Mlodzik et al., 1985). It was then shown that cad functions like a *Hox* gene product in specification of the posterior embryo as well as A-P patterning, particularly in the determination of the identity of the fly's most posterior abdominal segment (A10) (Moreno and Morata, 1999). *cad* is outside of the *Hox* cluster and its Hox-like function in A-P patterning in *Drosophila* has been explained by a similar evolutionary origin from a common ancestor named the *ProtoHox* gene cluster (Brooke et al., 1998). It is believed that the *ProtoHox* gene cluster duplicated in evolution to give rise to the *ParaHox* (initially discovered in the cephalochordate amphioxus and composed of *Cdx*, *Xlox* and *Gsx*) and the *Hox* gene

cluster (Brooke et al., 1998). It has also been proposed that, during vertebrate evolution, the *Hox* and *ParaHox* gene clusters duplicated to yield four mammalian *Hox* clusters (*Hoxa*, *Hoxb*, *Hoxc* and *Hoxd*) that give rise to 39 *Hox* genes and four *ParaHox* remnants loci, where a common *Cdx* member diverged to give rise to three *Cdx* genes (*Cdx1*, *Cdx2* and *Cdx4* in mouse) (Pollard and Holland, 2000). As it will be described below, this evolutionary divergence relocated vertebrate *Cdx* genes in a different strategic position, compared to *cad*, in the control of many processes of posterior development and A-P patterning.

0.6.2 The vertebrate *Cdx* genes

The vertebrate *Caudal* homologues form a multigene family composed of three *Cdx* genes (*Cdx1*, *Cdx2*, *Cdx4* in mammals; *CdxA*, *CdxB* and *CdxC* in chick; *Xcad1*, *Xcad2* and *Xcad3* in *Xenopus*; *cdx1a*, *cdx1b* and *cdx4* in zebrafish), with overlapping and dynamic caudal expression patterns (Blumberg et al., 1991; Bonner et al., 1995; Davidson and Zon, 2006; Duprey et al., 1988; Gamer and Wright, 1993; Gaunt et al., 2003; James and Kazenwadel, 1991; Lafontaine et al., 2012; Mallo et al., 1997; Marom et al., 1997; Reece-Hoyes et al., 2002).

In mammals, *Cdx* members are located on different chromosomes. In the mouse, *Cdx2* is located on chromosome 5 and is the only *Cdx* paralog that conserved its localization within the *ParaHox* cluster, together with *Pdx1 and Gsh2* (vertebrate homologs of *Xlox* and *Gsx*) (Chawengsaksophak and Beck, 1996; Pollard and Holland, 2000); *Cdx1* is located on chromosome 18 (Duprey et al., 1988), whereas murine *Cdx4* is X-linked (Horn and Ashworth, 1995). In human *CDX1*, *CDX2* and *CDX4* are located on chromosome 5q31-33, 13q12-13 and X, respectively (Bonner et al., 1995; Drummond et al., 1997; Horn and Ashworth, 1995).

Of the three mouse *Cdx* paralogs, *Cdx1* was the first to be discovered by Peter Gruss' lab in 1988, in an screen for mouse homologs of *Drosophila caudal* using a mouse

embryonic cDNA library and a *Drosophila caudal* gene probe (Duprey et al., 1988). *Cdx2* was subsequently identified in 1991 by Robert James and Jan Kazenwadel in a screen for homeobox-containing genes in the mouse adult intestine (James and Kazenwadel, 1991). The last one, *Cdx4*, was identified in 1993 by Laura W. Gamer and Christopher V.E. Wright from a mouse e8.5 cDNA library using an *XlHbox 8* probe bearing the 3' homeobox sequence (Gamer and Wright, 1993). There is no *Cdx3* gene, since it was initially identified in the pancreas of Syrian hamster but further analysis demonstrated that it is identical to mouse *Cdx2* (German et al., 1994).

0.6.2.1 Cdx spatio-temporal expression patterns in the mouse embryo

In the mouse embryo proper, *Cdx* expression starts with *Cdx1* expression at e7.5 in ectodermal and mesodermal cells of the PS (Meyer and Gruss, 1993). At around e7.75, before neurulation commences, *Cdx1* expression in the neuroectoderm reaches the level of the preotic sulcus in the prospective hindbrain region, whereas expression in the mesoderm is found slightly more posterior. At the onset of neurulation (e8.5), *Cdx1* is expressed in a gradient anterior (low) to posterior (high), with an anterior limit of expression at the level of the hindbrain/spinal cord boundary (Meyer and Gruss, 1993). During the wave of neurulation neural *Cdx1* expression regresses caudally to the spinal cord level at e9.5, and has been detected in trunk NCC that take the dorsolateral pathway (Meyer and Gruss, 1993). *Cdx1* expression in the caudal NT has been detected until e10.5. Of all *Cdx genes*, *Cdx1* is the only one expressed in the developing somites at e8.5, and then in the dermamyotome compartment of the anterior-most somites at e10.5 (Meyer and Gruss, 1993). *Cdx1* expression has also been detected in the developing limb buds from e9.5 to e10.5, as well as in the nephric cord and mesonephric ducts that eventually connect to the cloaca from e9.5 to e12 (Meyer and Gruss, 1993). *Cdx1* expression continues in the tailbud until extinction at e12. *Cdx1* transcripts are then detected in the epithelial lining of the intestine from e14 to adulthood (Duprey et al., 1988).

Cdx2 is the only *Cdx* member expressed in the extra-embryonic ectoderm tissue (ExE) during early embryonic development (Beck et al., 1995). Before implantation, at the blastocyst stage (e3.5), *Cdx2* is strongly expressed in the TE, where it plays a critical role in the specification of TE fate and embryo implantation (Strumpf et al., 2005). At post-implantation stage (e5.5), *Cdx2* expression is found in the ExE that derives from the TE (Beck et al., 1995). As the ExE gives rise to the ectoplacental cone and the chorionic ectoderm, *Cdx2* transcripts are confined to these tissues at mid-streak stage (e7.5) and then localize to the placenta until e12.5 (Beck et al., 1995). At e7.5, *Cdx2* transcripts are also found in the allantoic bud derived from the extra-embryonic mesoderm. In the embryo proper, *Cdx2* expression follows *Cdx1* and is strongly detected in all tissues of the PS at e8.5 as well as at the base of the allantois (Beck et al., 1995). Unlike *Cdx1*, *Cdx2* is not expressed in the somites and its mesodermal expression is limited to the unsegmented paraxial mesoderm. *Cdx2* transcripts are also expressed in a posterior high to anterior low gradient along the posterior embryo. Expression of *Cdx2* extends more anteriorly than *Cdx1* in the neuroectoderm with a limit to the somite 6 (Beck et al., 1995; Gamer and Wright, 1993). *Cdx2* is also expressed in the hindgut and cloacal endoderm. *Cdx2* expression continues in the tailbud until e12.5 (Beck et al., 1995). From this stage onwards, *Cdx2* expression is confined to the intestinal epithelium where it persist to adulthood (Beck, 2002).

Cdx4 is the last of the *Cdx* genes to be fully activated in the embryo proper and the first to extinguish its expression (Gamer and Wright, 1993). Expression starts at around e7.5, with low transcript levels being detected in the allantois and posterior tip of the PS. By early neurulation, high levels of *Cdx4* transcripts are detected in the three layers of the posterior half of the PS. At e8.5, *Cdx4* is expressed in the posterior neuroectoderm, presomitic mesoderm, lateral plate mesoderm and hindgut endoderm, with an anterior limit of expression located more posteriorly with respect to its paralogs. *Cdx4* expression in the paraxial mesoderm reaches the level of the most

recently formed somite and extends in the neuroectoderm a little more anteriorly to the level of the last 2 formed somites (Gamer and Wright, 1993). *Cdx4* is not expressed in the somites. By e9.5 *Cdx4* transcripts are confined to the tailbud where it extinguishes by e10.5. Like its paralogs, at all stages *Cdx4* exhibit a graded high posterior to low anterior expression pattern (Gamer and Wright, 1993).

In summary, in the mouse, all three *Cdx* genes are co-expressed in the caudal embryo at the onset of neurulation (e8.5), with a posterior (high) to anterior (low) gradient of transcript and protein distribution (Gaunt et al., 2003; Gaunt et al., 2005; Meyer and Gruss, 1993). The anterior boundaries of expression differ between *Cdx* members, *Cdx1* being the most anterior, followed by *Cdx2* and then *Cdx4*. The anterior limit of *Cdx*1 expression at E8.5 reaches the level of the hindbrain/spinal cord boundary. Of all *Cdx genes*, *Cdx1* is the only member expressed in the developing somites at E8.5 and then in the somite dermamyotome compartment. The three *Cdx* genes are expressed in the presomitic mesoderm. As the embryo elongates, *Cdx1*, *Cdx2* and *Cdx4* expression regresses caudally to the level of the spinal cord at E9.5 and is maintained in the tailbud until e10.5 for *Cdx4* and around e12 for *Cdx1* and *Cdx2* (Beck et al., 1995; Gamer and Wright, 1993; Meyer and Gruss, 1993). Then *Cdx1* and *Cdx2* are expressed in the hindgut endoderm with expression maintained postnatally (Beck, 2002). Regarding vital *Cdx* expression in extra-embryonic tissues, Cdx2 is expressed early at the blastocyst stage (e3.5) in the TE and is key for embryo implantation (Strumpf et al., 2005). The expression patterns of *Cdx1*, *Cdx2* and *Cdx4* in the mouse embryo from e7.5 to e9.5 are represented in Figure 0.14, (Lohnes 2003).

0.6.2.2 Regulation of murine *Cdx* expression. Cdx are downstream of the major signaling pathways

Studies in several vertebrate species have shown that *Cdx* expression is regulated by posteriorizing signals from canonical Wnt, RA and FGF pathways (Lohnes, 2003). BMP signaling has also been implicated in activation of *Cdx* expression in zebrafish,

Xenopus, mouse and human (Lengerke et al., 2008). Studies performed in mouse ES cells and zebrafish embryos suggest that BMP need to cooperate with canonical Wnt signaling to activate *Cdx* expression (Lengerke et al., 2008), whereas in human gastric cell lines the BMP-Smad pathway directly regulates *Cdx2* expression (Barros et al., 2008).

Of the posteriorizing signals, canonical Wnt and RA play essential and direct roles in the regulation of *Cdx* expression in the mouse (Houle et al., 2000; Houle et al., 2003; Lickert et al., 2000; Lickert and Kemler, 2002; Pilon et al., 2006; Pilon et al., 2007; Prinos et al., 2001; Sanchez-Ferras et al., 2012; Zhao et al., 2014). FGF signaling is required for *Xenopus* and chick *Cdx* expression (Bel-Vialar et al., 2002; Lengerke et al., 2008; Lohnes, 2003). However, as far as we know it is yet not clear whether FGF signaling is involved in mouse *Cdx* regulation. Indeed, perturbation of FGF signaling as seen in FGFR1 hypomorphs mouse mutants leads to vertebral transformations and consequently altered *Hox* expression but does not considerably affect *Cdx* expression (Partanen et al., 1998).

David Lohnes' lab has extensively characterized the mechanisms of regulation of mouse *Cdx1* expression, and shown that: 1) *Cdx1* is a direct target of RA and canonical Wnt signals *in vivo*; 2) RA and Wnt act synergistically in the activation of *Cdx1* expression at e7.5; and 3) at later stages (by e8.5-e9.5) canonical Wnt and Cdx1 itself maintains its own expression via a Wnt-ß-catenin-dependent auto-regulatory loop (Beland et al., 2004; Houle et al., 2000; Houle et al., 2003; Pilon et al., 2007; Prinos et al., 2001). First, Lohnes' group identified one atypical RA response element (RARE) in the *Cdx1* proximal promoter (Houle et al., 2000). Then, by generating a *Cdx1*RARE-null mouse mutant line, the authors showed a critical role for this regulatory element in the regulation of *Cdx1* expression and function *in vivo* (Houle et al., 2003). *Cdx1*RARE-null mutants exhibited a reduced *Cdx1* expression at e7.5-e8.5, leading to homeotic vertebra transformations characteristics of *Cdx1*

heterozygous mutants (Houle et al., 2003). The mild effect of the RARE mutation on *Cdx1* expression suggested that RA signal is required to tune *Cdx* expression levels and that other signal are necessary to induce *Cdx* expression.

By that time, David Lohnes' and other labs were also characterizing the regulation of *Cdx1* by canonical Wnt and identified two functional Lef/Tcf response elements (LRE1 and LRE2) in the *Cdx1* proximal promoter (Lickert et al., 2000; Lickert and Kemler, 2002; Prinos et al., 2001). It was shown that these LREs are essential for the response of the *Cdx1* promoter to canonical Wnt activation *in vitro* and for reporter gene expression *in vivo* (Lickert and Kemler, 2002; Prinos et al., 2001). *In vitro* assays performed in the F9 embryocarcinoma cell line also showed that canonical Wnt and RA signaling synergistically activate the *Cdx1* proximal promoter and that the two LREs and the RARE mediate this effect (Prinos et al., 2001). Besides, this work provided the first *in vivo* evidence that *Cdx1* autoregulates, by showing that *Cdx1* expression is reduced in e8.5 *Cdx1* null embryos and that Cdx1 can activate expression from its own promoter (Prinos et al., 2001).

The importance of the LRE elements and canonical Wnt for endogenous *Cdx1* expression *in vivo* was then studied by the generation of mouse mutants for the two LREs or LREs + RARE in the endogenous *Cdx1* proximal promoter (Pilon et al., 2007). LRE mutant embryos phenocopied the *Cdx1*- null mutants, demonstrated by a complete failure in induction of *Cdx1* expression and the presence of *Cdx1* null-like homeotic transformations (Pilon et al., 2007). The LRE-mutated *Cdx1* promoter was less responsive to RA. Taken together, this work showed a critical role for canonical Wnt signaling in *Cdx1* induction *in vivo* as well as in mediating the activation of the *Cdx1* promoter by RA. In parallel to this work was the demonstration that canonical Wnt also directly regulates *Cdx4* expression *in vivo* and the identification of functional LREs in the *Cdx4* 5'-proximal promoter mediating Lef1/ßcatenin binding *in vivo* and transcriptional response in embryocarcinoma cells (Pilon et al., 2006).

Other studies including this work have shown that *Cdx2* is also a direct target of canonical Wnt (Sanchez-Ferras et al., 2012; Zhao et al., 2014).

From e8.5 onwards, RA signaling is excluded from the tailbud (Sakai et al., 2001). RA activity is restricted to more anterior regions of the trunk, where *Cdx genes* are not expressed. This has two main implications. First, it is unlikely that the RA signal is implicated in the activation of *Cdx2* and *Cdx4* since their expression in the embryo proper begins at e8.5 and very probably because of that there are no reports of RA signaling regulating *Cdx2* and *Cdx4.* Hence, canonical Wnt may be the major signaling pathway involved in *Cdx2* and *Cdx4* activation (Pilon et al., 2006). Second, other mechanisms must impede RA activation of *Cdx* expression at more anterior levels. In this regard, work from David Lohnes' lab proposed a mechanism in which RA induces expression of the transcriptional repressors chicken ovalbumin upstream promoter-transcription factors (COUP-TFs) which then binds to the RARE and prevents activation of the *Cdx1* promoter by RA (Beland and Lohnes, 2005).

0.6.2.3 Cdx autoregulation and cross-regulatory interactions

Cdx expression is not only regulated by the major signaling pathways that orchestrate embryo development but also by auto and cross-regulatory loops. Work from David Lohnes' lab has also contributed to the understanding of the mechanisms underlying these regulations (Beland et al., 2004; Grainger et al., 2013; Prinos et al., 2001; Savory et al., 2011b). As described above, this group first found that *Cdx1* is required to maintain its own expression *in vivo* from e8.5 onwards, and that Cdx1 transactivates its proximal promoter in co-transfection analysis performed in the P19 embryocarcinoma cell line (Prinos et al., 2001). Further characterization of this autoregulation revealed that Cdx1 is tethered to its own promoter via direct physical interaction with the Lef1-ß-catenin complex which binds to the LREs. This allows Cdx1 to transactivate its expression without direct DNA binding (Beland et al., 2004). Protein-protein binding assays showed that Cdx1 directly interact via its

homeodomain with the high mobility group (HMG) box of Lef1 (Beland et al., 2004). Given the high conservation of these DNA binding domains between members of the Cdx and Lef/TCF families, this interaction has also been found for Cdx1-TCF4, Cdx2-Lef1, and Cdx2-TCF4 (Grainger et al., 2013).

The next obvious question is whether Cdx2 or Cdx4 can regulate *Cdx1* expression or vice-versa. This has been addressed by a gene replacement strategy, in particular by generating a knock-in mouse model in which *Cdx1* is replaced by *Cdx2* (called $Cdx1^{2ki/2ki}$). *Cdx2* open reading frame was targeted into the *Cdx1* locus, leading to Cdx1 inactivation and expression of *Cdx2* under the control of the *Cdx1* promoter (Savory et al., 2009b). Analysis of $Cdx1^{2ki/2ki}$ has revealed that Cdx2 can support autoregulation of the *Cdx1* locus during vertebral patterning (Savory et al., 2009b) but not in the intestine (Grainger et al., 2013), thus suggesting a cell-type dependent function for Cdx2 in regulation of the *Cdx1* promoter. Cdx1 and Cdx2 specific functions in the intestine have been attributed to differences in the N-terminal sequences, and probably to the specific interaction of the Cdx1 transactivation domain with an intestinal co-factor (Grainger et al., 2013).

In some particular context, such as in human colon cancer cell lines, Cdx2 inhibits activation of *Cdx1* by canonical Wnt signaling (Domon-Dell and Freund, 2002). This has been explained by the finding that Cdx2 binds ß-catenin and dissociates the ß-catenin-TCF complex (Guo et al., 2010). On the other hand, Cdx2 can regulate its own expression in the intestine and this function is also cell type specific (Xu et al., 1999). It has been shown that Cdx2 directly regulate its own 1.2 kb proximal promoter in human gastric cancer cell lines via binding to three Cdx binding sites (CdxBS), with ChIP analysis from samples of the mouse ileum and human gastric intestinal metaplasia showing Cdx2 occupancy of its own proximal promoter *in vivo* (Barros et al., 2011). Cdx2 is also a direct regulator of *Cdx4* expression *in vivo*, and fulfils this function via direct binding to the DNA, without requirement of canonical

Wnt input (Savory et al., 2011b). Taken together, this work illustrates that Cdx auto- and cross-regulatory mechanisms are complex and context-specific.

0.6.2.4 Interactions between Cdx and canonical Wnt

Cdx genes are not only targets of canonical Wnt signaling but their gene products can also modulate this pathway at least via two mechanisms. First, Cdx proteins can activate the expression of canonical Wnt ligands. Indeed, studies in *Xenopus* and mice have evidenced the existence of a Wnt3a-Cdx positive feedback loop (Faas and Isaacs, 2009; Savory et al., 2009a). In the mouse, it has also been shown that Cdx2 directly activate the expression of Wnt3a, and maintains Wnt activity in the paraxial mesoderm at E8.5 (Savory et al., 2009a). On the other hand, Cdx proteins can directly physically interact with the nuclear effector of canonical Wnt signaling, the Lef/Tcf-ß-catenin transcriptional complex to activate expression of Wnt target genes (Beland et al., 2004). Hence, Cdx can activate transcription in a Wnt-ßcatenin dependent manner via LREs independently of the presence of Cdx binding sites in target promoters.

0.6.2.5 Cdx protein structure. Structural basis for functional overlap and specificity between Cdx members

Vertebrate Cdx protein structure is characterized by the presence of: 1) a short N-terminal sequence of approximately 10 amino acids, which has been implicated in cytoplasmic export of the protein (Trinh et al., 1999); 2) a potent N-terminal transactivation domain which has been mapped to position 60-140 in the amino acid sequence of Cdx2 (Rings et al., 2001) and 3) a DNA-binding homeodomain of 180 bp long (60 amino acids) towards the C-terminus (Beland et al., 2004; Duprey et al., 1988; Gamer and Wright, 1993; Lohnes, 2003; Marom et al., 1997; Rings et al., 2001). Protein size varies between murine Cdx members, with Cdx1 exhibiting 28 kDa (268 amino acids, P18111 UniProtKB); Cdx2 33 kDa (311 amino acids, P43241 UniProtKB) and Cdx4 30kDa (282 amino acids, Q07424 UniProtKB). Differences in

size between Cdx proteins have been associated with the relative low sequence homology in the N-terminal transactivation domain. This region is also subject to post-translational modifications (Marom et al., 1997; Rings et al., 2001). In particular, it has been shown that the Cdx2 activation domain is phosphorylated at serine 60 by the MAPK pathway (Rings et al., 2001).

Sequence alignment studies have revealed a high degree of homology in the homeodomain region between Cdx members and other Caudal-related proteins (Figure 0.14) (Gamer and Wright, 1993; Marom et al., 1997). For example, the homeodomain of mouse Cdx1, 2 and 4 exhibit around 90% sequence identity (Gamer and Wright, 1993); 84% between mouse Cdx1 and *Drosophila* cad (Marom et al., 1997) as well as 100% between mouse and human Cdx1 (Mallo et al., 1997). The high degree of sequence conservation of the homeodomain has been proposed to account for the similar functions of Caudal-related proteins in the control of posterior development in Bilaterians, and functional overlap of Cdx members (Lohnes, 2003). Other regions of high homology include the N-terminal-10 amino acids sequence (A domain); a 9 amino acid stretch upstream of the homeodomain (B domain), a hexapeptide motif located upstream of the homeodomain (C domain) and the amino acids flanking the homeodomain (Gamer and Wright, 1993; Marom et al., 1997).

Outside the homeodomain, there is limited sequence homology in the Caudal-related family. Of note is the low sequence similarity in the N-terminal transactivation domain of murine Cdx proteins (Marom et al., 1997). This has been proposed to account for some specific functions of Cdx members, for example, the unique role of Cdx2 in the control of intestinal endoderm patterning (Grainger et al., 2010), as well as the different transactivating activity of Cdx proteins in tissue culture (see (Grainger et al., 2013) and this study). In addition, it has been shown that Cdx2 transcriptional activity is modulated by MAPK-dependent phosphorylation of serine

60 residue in the activation domain (Rings et al., 2001), thus providing another explanation for the differential activity of Cdx proteins.

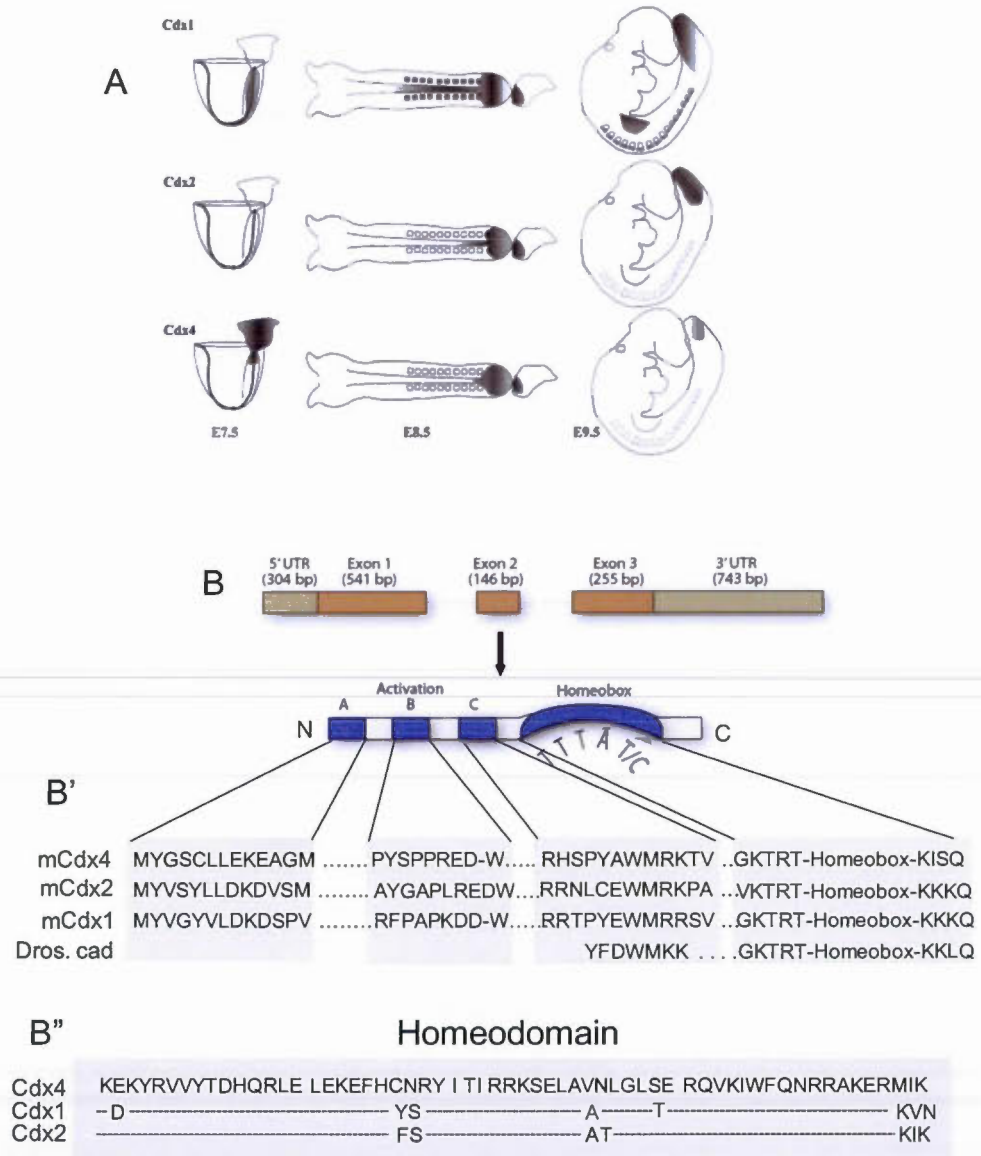


Figure 0.14 Comparison of expression pattern, genomic and protein structure of mouse Cdx members.

(A) Schematic representation of Cdx1, 2 and 4 expression patterns in e7.5, e8.5 and e9.5 mouse embryos. Source of the image: Lohnes, 2003. (B) Schematic representation of the genomic and protein structure of the human CDX2 gene. Domains of high sequence

homology are represented in blue and include motifs A, B, and C located within the N-terminal activation domain and the homeobox binding domain. Adapted from (Witek, 2012), transcription factor encyclopedia. B' and B": Comparison of amino acid sequence between conserved domains of murine (m) Cdx1, Cdx2, Cdx4 and *Drosophila* (Dros) cad proteins. Modified from Gamer and Wright, 1993.

0.6.2.6 Mechanisms of Cdx transactivating action

As homeodomain transcription factors, Cdx are expected to regulate gene transcription via binding to AT rich sequences within regulatory regions of their target genes. The consensus Cdx binding site: A/CTTTATA/G, was defined many years ago based on in vitro binding studies of the chicken CdxA on random oligonucleotides and genomic DNA fragments (Margalit et al., 1993). The Cdx consensus binding site is very similar to that of *Drosophila* cad: TTTATG (Dearolf et al., 1989); consistent with the high sequence similarity in the homeodomain between these evolutionarily distant Caudal-related members.

Following these binding site studies, a considerable number of Cdx targets have been reported (Beland et al., 2004; Charite et al., 1998; Chun et al., 2007; Grainger et al., 2011; Savory et al., 2009a; Savory et al., 2011a; Savory et al., 2011b; Sturgeon et al., 2011; Subramanian et al., 1995). In most cases, Cdx proteins regulate gene expression by directly binding to consensus-like response elements. These studies have also shown that Cdx may act either as activators or repressors of target gene expression, although there are very few examples for the latter. Only two such Cdx negative targets have been reported: the hindbrain patterning gene *Mafb*, which is directly and transiently repressed by Cdx1 posterior to rhombomeres 6/7 boundary (Sturgeon et al., 2011) and the *insulin-like growth-factor-binding protein 3* gene, which is repressed by Cdx2 in the human colonic adenocarcinoma cell line LOVO (Chun et al., 2007).

The molecular mechanisms by which Cdx proteins regulate gene transcription are poorly understood. The implication of Cdx in either activation or repression likely

implies interactions with co-activators or co-repressor proteins. Also, whereas direct binding to cognate sites is a general way of action of Cdx proteins, these homeoproteins have also been shown to influence gene transcription via interaction with co-activators or specific nuclear cofactors, without necessarily binding to the DNA. One of the best examples is the ability of Cdx1 to activate its own expression via interaction with the Lef1/TCF-ß-catenin complex on Lef-TCF binding sites (Beland et al., 2004).

A well known co-activator and direct interacting partner of Cdx is the cAMP response element-binding protein (CREBP)-binding protein CBP, which has been shown to potentiate Cdx2 trans activity in the Caco-2 colonic cell line (Lorentz et al., 1999). Cdx2 has also been shown to directly physically interact with the closely related CBP homologue and co-activator p300 and, together with Pax6, form a transcriptional complex that potentiates activation of the *glucagon* promoter (Hussain and Habener, 1999). Other, tissue specific, described Cdx interacting partners include Sox2, Brn-4, Pbx1, TCF4, GATA6, HNF4A, HNF1α, pSmad1/5/8 as well as a subset of Hox homeoproteins including Hoxd4, Hoxa6 and Hoxa9 (Beland et al., 2004; Lafontaine et al., 2012; Liu et al., 2006b; Mari et al., 2014; Mitchelmore et al., 2000; Verzi et al., 2010a; Verzi et al., 2010b; Wang et al., 2006).

Intriguingly, Cdx proteins appear to poorly exploit the hexapeptide motif which is normally used for protein-protein interaction (Chang et al., 1995), since in most cases interaction has been shown to be mediated by the homeodomain (Beland et al., 2004; Lafontaine et al., 2012; Mitchelmore et al., 2000; Wang et al., 2006). The N-terminal region has been implicated in transcriptional activity. The molecular basis of this function is unknown, although probably involves the recruitment of co-regulators. As described above there is low conservation in this region between Cdx members, and interaction with specific co-regulators via the N-terminal domain may contribute to

the differences in trans activity as well as the functional specificity observed in intestinal tissues (Grainger et al., 2013).

Clusters of CdxBS are usually found within the proximal promoter of target genes. Regardless of the enhancer activity of these regions, Cdx may be important components of the basal transcriptional machinery, which tether transcriptional co-activators and stabilize the transcription initiation complex. Indeed, such a role has been demonstrated for Cdx proteins in the activation of intestinal epithelial cell gene promoters via recruitment and interaction with the co-activators CBP/p300, as previously mentionned (Hussain and Habener, 1999; Lorentz et al., 1999). A recent study shed lights on the mechanism of core promoter-specific transcriptional regulation by *Caudal* genes (Shir-Shapira et al., 2015). Shir-Shapira and co-workers showed that the C-terminus domain of *Drosophila* Caudal and the N-termini of murine Cdx proteins, as well as the conserved homeodomain, are important for core promoter activation and that CBP acts as a Caudal co-activator that mediates interaction between Caudal binding at the enhancer and the core promoter region. Both CBP and p300 are proposed to serve as adapter proteins that recruit transcription factors as well as the RNA polymerase II and assemble the transcription initiation complex (Hussain and Habener, 1999; Ogryzko et al., 1996). Via interaction with CBP and p300 Cdx may also modulate access to the DNA since these co-activators possess histone acetyltransferase activity that enhances access to the DNA (Ogryzko et al., 1996).

0.6.2.7 Cdx functions in embryogenesis

At the difference of *Cad* that acts as a *Hox*-like gene in the specification of posterior identity (Moreno and Morata, 1999), vertebrate *Cdx* gene products act upstream of *Hox* genes and pattern the mesoderm and neuroectoderm along the A-P axis by transmitting positional identity from the posteriorizing canonical Wnt, RA and FGF pathways to *Hox* promoters (Bel-Vialar et al., 2002; Copf et al., 2004; Houle et al.,

2000; Houle et al., 2003; Keenan et al., 2006; Lengerke et al., 2008; Lohnes, 2003; van den Akker et al., 2002). Therefore, it appears that given the importance of these signals in development of the three germ layers in vertebrates, evolution co-opted *Cdx* to new functions and gave *Cdx* a new hierarchical position in the control of posterior morphogenesis.

In the recent years, the use of conditional mutagenesis approaches allowed to circumvent the peri-implantation lethality of *Cdx2*-null embryos (Strumpf et al., 2005), and revealed several important *Hox*-dependent and *Hox*-independent roles for Cdx proteins. Analysis of allelic series of *Cdx* compound mouse mutants showed that Cdx proteins act redundantly and in a *Hox*-dependent manner in the control of A-P patterning, hematopoiesis and urorectal septation (van de Ven et al., 2011; van den Akker et al., 2002; Wang et al., 2008b). Importantly, Cdx also impact on axial elongation, somitogenesis, intestinal endoderm development, placentogenesis and NT closure independently of *Hox* genes (Gao et al., 2009; Grainger et al., 2010; Savory et al., 2009a; Savory et al., 2011a; van Nes et al., 2006; Young et al., 2009). One particular caveat to these models is that Cdx function has been inactivated in all germ layers and is not informative as whether Cdx roles in these processes is tissue-autonomous or not. Cdx functional overlap as well as the severe NT closure and posterior truncation phenotypes of *Cdx* compound mutants (Savory et al., 2011a; van Rooijen et al., 2012) are expected to mask new, yet unknown, Cdx roles.

In adulthood, Cdx1 and Cdx2 are master regulators of intestinal epithelial homeostasis and have been implicated in the control of cell cycle exit, proliferation, apoptosis, differentiation, cell-adhesion and morphology by regulating expression of a considerable number of genes including *Cyclin D1/D2, p21*$^{WAF/CIP}$ and *Bcl-2*. By these and other mechanisms Cdx1 and Cdx2 have also been implicated in colon carcinogenesis. Particularly, *Cdx1* and *Cdx2* expression is turned-off in a subset of human colorectal cancer (CRC) and is a bad prognostic of progression of this life

threatening disease (Baba et al., 2009; Pilozzi et al., 2004). Recent work from David Lohnes' lab clearly evidenced a role for Cdx1 and Cdx2 as supressors of CRC in murine models, by regulating TGF-ß and Eph-Ephrin signaling (Hryniuk et al., 2014). However, other studies have found that *Cdx2* is overexpressed in certain CRC where it acts as a lineage-survival oncogene in cooperation with Wnt/ß-catenin signaling (Salari et al., 2012). Taken together, these studies suggest that Cdx proteins role in CRC is complex and poorly understood. For excellent review on adult functions of Cdx see Guo et al., 2004.

Hox-dependent roles for Cdx proteins in the control of A-P vertebral patterning

Perhaps one of the most well established Cdx roles in the mesoderm is the regulation of A-P vertebral patterning (Charite et al., 1998; Lohnes, 2003; Subramanian et al., 1995; van den Akker et al., 2002; van Nes et al., 2006). The analyses of allelic combinations of *Cdx* mutant mice clearly showed that Cdx members redundantly control this process. *Cdx* loss-of-function results in anterior homeotic transformations of the vertebral skeleton that increase in severity in $Cdx1^{-/-}$::$Cdx2^{+/-}$ compound mutants, compared to single $Cdx1^{-/-}$ and $Cdx2^{+/-}$ mutants (Table I) (van den Akker et al., 2002). Cdx4 synergizes with Cdx1 and Cdx2 in the control of vertebral A-P patterning, although its contribution to this process appears to be minor (van Nes et al., 2006).

Analyses of *Cdx* gain- and loss-of-function models in mice and other vertebrate species have revealed that Cdx proteins regulate the *Hox* code from cervical to caudal axial levels (see (Lohnes, 2003) for review); (Bel-Vialar et al., 2002; Charite et al., 1998; Houle et al., 2003; Isaacs et al., 1998; Subramanian et al., 1995; van den Akker et al., 2002). Consistent with this, *Cdx* loss-of-function leads to a posterior shift in the anterior boundary of expression of *Hox* genes. For example, e9.5 $Cdx1^{-/-}$ mutant embryos exhibit a posterior shift in the anterior mesodermal limit of the *Hox* paralog group 3 (*Hoxa3*, *Hoxb3* and *Hoxd3*) from somite 5 to 6, which correlates with the

observation of anterior homeotic transformation in the first and second cervical vertebrae in *Cdx1* $^{-/-}$ offspring (Houle et al., 2003; Subramanian et al., 1995). On the other hand, it has been shown that *Cdx* gain-of-function causes an anterior shift in *Hox* expression boundary, thus pointing to the importance of Cdx dosage in the regulation of *Hox* expression (Bel-Vialar et al., 2002; Charite et al., 1998). However, later work from David Lohnes' lab failed to see defects in vertebral patterning and anterior shift in *Hox* expression in a BAC transgenic mouse line overexpressing Cdx1, thus suggesting that a certain threshold of Cdx overexpression must be attained in order to impact in vertebral patterning and *Hox* expression (Savory et al., 2009b).

Clusters of potential CdxBS have been found in the promoter region of a considerable number of mouse and human *Hox* genes (Subramanian et al., 1995), suggesting that several *Hox* genes are direct Cdx targets. Consistent with this, *in vitro* and transgenic reporter analyses have shown that intact CdxBS or the appropriate number of CdxBS are required for the proper expression and spatial restriction of *Hoxa7* and *Hoxb8* (Charite et al., 1998; Gaunt et al., 2004; Subramanian et al., 1995). Whereas Cdx proteins are clearly involved in the regulation of a number of *Hox* genes, it is not fully understood how the promoter of these genes respond to a caudal *Cdx* gradient. One of the hypothesis suggest a differential sensitivity to Cdx dosage along the A-P axis (Houle et al., 2003), which at the cis-regulatory level has been proposed to be mediated by the number or affinity of CdxBS (Lohnes, 2003). This hypothesis is based on the observation that increasing copy number of CdxBS by multimerization of *Hoxa7* and *Hoxb8* enhancers results in anteriorization of transgenic reporter expression in the mesoderm and neuroectoderm (Charite et al., 1998; Gaunt et al., 2004).

Transplantation experiments performed in chick embryos indicate that the *Hox* code is fixed in paraxial mesoderm cells from the PSM region before somites form (Kieny et al., 1972). This zone is exposed to counteracting gradients of Wnt/FGF and RA

morphogens that act to establish the Hox code (Deschamps and van Nes, 2005). The spatiotemporal Cdx expression pattern is consistent with a role in conveying posteriorizing signals in the PSM. Indeed, although loss of one allele of *Cdx2* affects the patterning of cervical vertebra, *Cdx2* mesodermal expression is restricted to the PSM region (Lohnes, 2003).

A role for Cdx as intermediates in the mechanism of regulation of certain *Hox* genes by posteriorizing canonical Wnt, RA and FGF signaling is supported by several studies. One of the best examples is that in mice, inactivation of the response of the *Cdx1* promoter to canonical Wnt and RA signaling by mutation of LREs and RARE results in posterior shifts of expression of *Hox* genes from the groups 3 and 4, indicating that posteriorizing signals act via regulation of Cdx1 to control *Hox* expression (Pilon et al., 2007). However, the existence of Cdx-independent pathways including a direct regulation of *Hox* genes by posteriorizing signals including RA has also been reported (Lohnes, 2003). In support of this, the *Cdx* loss-of-function models generated so far have shown mild changes in anterior expression boundaries of *Hox* genes. Triple *Cdx*-null mutants completely lack posterior tissues which impairs the analysis of the effect of pan-Cdx loss-of-function in trunk *Hox* gene expression (van Rooijen et al., 2012). This suggests the necessity of using conditional approaches that circumvent posterior truncation and embryonic lethality of pan-Cdx loss-of-function models.

Hox-independent roles of Cdx in the control of axial elongation

Mouse embryos that completely lack Cdx1, Cdx2 and Cdx4 function due to inactivation in the epiblast fail to generate trunk and tail structures and only have a head and occipital tissues (van Rooijen et al., 2012). This and previous studies have undoubtedly evidenced the essential Cdx role in development of the posterior embryo (Chawengsaksophak et al., 2004; Savory et al., 2009a; Savory et al., 2011a; van Rooijen et al., 2012). Such a role for *Caudal*-related genes has been evolutionarily

conserved in Bilaterian species that elongate their axis by adding trunk and tail tissues from a posterior growth zone (Copf et al., 2004).

Of the three Cdx members, Cdx2 has the most important contribution to axial elongation. Loss of one *Cdx2* allele results in the presence of a foreshortened tail, whereas *Cdx1*- or *Cdx4*-null mice, do not exhibit overt axial elongation phenotypes (Chawengsaksophak et al., 1997). However, the progressive increase in severity of axial truncation phenotypes in compound mutants clearly shows that Cdx members act redundantly in the control of axial elongation. For example, truncations are observed at the level of the hindlimbs in *Cdx2* conditional null embryos, (Chawengsaksophak et al., 2004; Savory et al., 2009a), rostral to the forelimbs in *Cdx1-Cdx2* double mutants (Savory et al., 2011a) and at post-occipital levels in triple *Cdx1-Cdx2-Cdx4* mutants (van Rooijen et al., 2012).

Analyses of compound *Cdx* mutants have revealed crucial redundant roles for Cdx members in the coordination of axial elongation, somitogenesis and A-P patterning of the somites (Savory et al., 2009a; Savory et al., 2011a; van Rooijen et al., 2012; Young et al., 2009). As mentioned previously, this is consistent with the expression of all three *Cdx* genes in the caudal embryo during the time these processes take place and the known Cdx role as integrators of signaling inputs from canonical Wnt, FGF and RA signaling. Studies from several groups suggest that Cdx function is required for the maintenance of the expression of a number of genes that encode important regulators of posterior axial growth, generation, maturation and segmentation of the PSM such as *Wnt3a*, *Fgf8*, *T*, *Tbx6*, *Cyp26a1*, the Notch ligand *Dll1*, *Msgn1*, as well as the somitic mesoderm markers *Mox1*, *Paraxis* and *Uncx4.1* (Chawengsaksophak et al., 2004; Grainger et al., 2012; Savory et al., 2009a; Savory et al., 2011a; van Rooijen et al., 2012; Young et al., 2009). Some of these genes, particularly *Wnt3a*, *T*, *Cyp26a1* and *Dll1*, have been validated as direct Cdx targets by promoter and chromatin immunoprecipitation analyses (Grainger et al., 2012; Savory et al., 2009a).

The effect of pan-Cdx loss-of-function on posterior development and *Wnt3a* and *T* gene expression manifest after 5 somites have been generated (around e8.5) and expression of these genes is correctly initiated. This is consistent with a model in which Cdx are initially downstream of the posteriorizing inputs and then act upstream of these signals (positive feedback loop on canonical Wnt) and Notch signal to maintain the transcriptional program that orchestrates trunk and tail development (Grainger et al., 2012; Savory et al., 2011a; van Rooijen et al., 2012). As a consequence of this crucial Cdx role, the posterior growth zone loses its progenitor activity at around e8.5, probably because of a failure in clearance of RA from this region, and no new tissue in all three germ layers is generated after this stage in *Cdx* null mutants (van Rooijen et al., 2012). It has been shown that axial elongation defects can be rescued by reactivation of either FGF or canonical Wnt signaling as well as by gain-of-function of trunk *Hox* genes (van Rooijen et al., 2012; Young et al., 2009). Taken together, these observations indicate that Cdx proteins use a battery of Hox-independent means as well as cooperate with central *Hox* genes to ensure posterior development.

Caudal neural functions

In chordates, *Caudal*-related genes are strongly expressed in the posterior neuroectoderm and, consistent with this, have been shown to play important roles in neural development. Studies performed in lower and higher chordate species suggest an evolutionarily conserved role for *Caudal*-related genes in the control of NT closure (Katsuyama et al., 1999; Mita and Fujiwara, 2007; Savory et al., 2011a). For example, in the urochordate ascidian embryo *Ciona Intestinalis*, neural specific Cdx loss-of-function, via expression of a dominant negative form of the Cdx homologue Ci-Cdx, impairs normal NT closure (Mita and Fujiwara, 2007). A similar strategy also evidenced an important role in NT closure in the ascidian species *Halocynthia roretzi* (Katsuyama et al., 1999). In mice, a recent study clearly demonstrated that Cdx function is required for NT closure along the whole spinal cord region (Savory et

al., 2011a) (described below). Other, yet unknown important neural roles for Cdx are expected to be masked by functional overlap of Cdx members as well as by the embryonic lethality and severe truncation phenotypes of triple Cdx mutants.

The spatiotemporal expression pattern in the caudal neurectoderm, and position downstream of posteriorizing signaling inputs, make Cdx excellent candidates for being involved in the coordination of NC ontogenesis. In support of this, the study performed in the ascidian embryo *Halocynthia roretzi* not only evidenced a Cdx role in NT closure and tail elongation but also in pigment cell development (Katsuyama et al., 1999). A recent study has shown that ascidian pigment cells constitute the ancestral homologs of melanocytes and represent a rudimentary NC (Abitua et al., 2012). Such a role for Cdx proteins in NC-derived pigment development has not yet been shown in vertebrates.

Cdx and the control of neural tube development in mouse

Analyses of *Cdx* compound mouse mutants show that these genes are essential for normal NT development (Savory et al., 2011a; van de Ven et al., 2011). *Cdx1/2* cDKO mutants exhibit craniorachischisis – the most severe NTD – that extends along the whole posterior axis (Savory et al., 2011a). Characterization of this phenotype suggests that Cdx proteins impact on NT closure by regulating the expression of members of the PCP pathway and consequently convergent extension movements, with the PCP receptor *Ptk7* being a direct Cdx target (Savory et al., 2011a). On the other hand, overexpression of Cdx members in the NT and somites can cause NTD (Charite et al., 1998). Altogether, these studies suggest that Cdx act redundantly and in a dose-dependent manner in the control of NT closure. *Cdx2/4* cDKO mutants exhibit caudal-specific defects in arrangement of the neurepithelium and misspecification of the mesoderm into ectopic neural tubes (van de Ven et al., 2011). The characterization of posterior NT morphology defects in these mutants suggests a role for Cdx proteins in the organization of the caudal neuroepithelium and apico-

basal polarity in a canonical Wnt and *Hox*-dependent manner (van de Ven et al., 2011). However, since the mutagenesis approaches used in these studies has involved inactivation of Cdx function in the three germ layers, it is currently unknown whether Cdx fulfill these important roles in an autonomous manner or indirectly via the mesoderm. Interestingly, in the ascidian embryo *Ciona Intestinalis,* the Cdx homologue Ci-Cdx directly controls NT closure by taking part in a regulatory cascade involving Nodal signaling and Zic homologues (Mita and Fujiwara, 2007). Whether Cdx operate by a similar mechanism in the mouse is currently unknown.

Cdx role in patterning of NT and NCC

Several studies have clearly implicated Cdx in the regulation of the neurectodermal expression of trunk *Hox* genes in vertebrates (Bel-Vialar et al., 2002; Charite et al., 1998; Deschamps and van Nes, 2005; Isaacs et al., 1998; Skromne et al., 2007). For example, in chick embryos, electroporation of a dominant negative form of *Xenopus* Cdx (*XcadEnR*) in the NT inhibits the upregulation of *Hoxb9* expression by FGF2 treatment (Bel-Vialar et al., 2002). Consistent with this, *in vitro* assays of chick neural cell differentiation suggest that Cdx proteins are involved downstream of posteriorizing signals and in a *Hox*-dependent manner in the A-P patterning of neural progenitors from the hindbrain and spinal cord into different motor neuron subtypes (Nordstrom et al., 2006). *Hox* genes are also implicated in rostrocaudal patterning of NC-derived spinal ganglia (Charite et al., 1994; van den Akker, 1999). Interestingly, $Cdx1^{-/-}$::$Cdx2^{+/-}$ double, but not single mouse mutants, exhibit spinal ganglia fusions at cervical and thoracic level, which correlate with a transient posterior shift in the anterior expression boundary in the neuroectoderm of certain 5' *Hox* genes like *Hoxb8* (van den Akker et al., 2002). This suggests that Cdx proteins act redundantly in the control of trunk NC patterning via regulation of *Hox* genes as well. However, this study did not make clear a direct role for Cdx in NC patterning, since vertebral processes were also found to be fused in $Cdx1^{-/-}$::$Cdx2^{+/-}$ mutants, and it is known that somite properties affect the migratory capacity of NCC (Bronner-Fraser and Stern,

1991). On the other hand, studies in zebrafish embryos have revealed that Cdx proteins act redundantly, in an autonomous manner and independently of *Hox* genes to control the correct positioning of the hindbrain-spinal cord boundary (Skromne et al., 2007). Particularly, gain- and loss-of-function as well as transplantation studies suggest that Cdx are required for the specification of the spinal cord identity via repression of the hindbrain developmental program. This work also reported that Cdx loss-of-function affects the number of Rohon-Beard mechanosensory neurons, which share a common precursor with NCC (Epperlein et al., 2007; Skromne et al., 2007). In summary, all these observations indicate that Cdx impacts on neural development via *Hox*-dependent and *Hox*-independent mechanisms.

0.7 Hypothesis and objectives

Cdx genes are strongly expressed in the neuroectoderm across chordates, but their function in this tissue has long remained obscure. We were particularly intrigued by their high expression in the neuroectoderm at the onset of NC induction, their strategic position downstream of the major signaling pathways that initiate the NC developmental program (i.e. canonical Wnt, FGF and BMP) (Betancur et al., 2010; Lengerke et al., 2008; Lohnes, 2003; Pilon et al., 2006; Pilon et al., 2007), as well as by the poor evidence of direct regulatory interactions between these pathways and the neural plate border specifier module composed by *Pax3/7*, *Msx1/2* and *Zic2/5* (Betancur et al., 2010; Stuhlmiller and Garcia-Castro, 2012). Based on these observations, we hypothesized that Cdx proteins are important novel regulators of trunk NC development in the mouse, and act at the top of the trunk NC-GRN to convey positional information from the signaling module to the neural plate border specifier module in order to start the NC developmental program.

Pioneer work in *Xenopus* embryos established a central role for posteriorizing signals from the canonical Wnt pathway in the induction of *Pax3* neural expression (Bang et al., 1999). We were particularly interested by this regulation because of the

importance of the canonical Wnt-*Pax3* circuit in NC induction (Barembaum and Bronner-Fraser, 2005; Garcia-Castro et al., 2002; LaBonne and Bronner-Fraser, 1998; Monsoro-Burq et al., 2005) and the multiple evidence of interactions between Cdx genes/proteins and canonical Wnt signaling (described below). When this thesis project began, around four years ago, it was unclear how positional information from canonical Wnt signaling is integrated at the cis-regulatory level to initiate *Pax3* expression at the mouse PNP border. Based on several elements, we reasoned that Cdx proteins could act as intermediates in this mechanism and relay positional information from the canonical Wnt pathway to the *Pax3* promoter: 1) *Pax3* was known as one of the first markers of pre-migratory NCC of the trunk and Cdx proteins are highly expressed in the caudal neuroectoderm at the onset of *Pax3* expression (Goulding et al., 1991; Meyer and Gruss, 1993); 2) Murine *Cdx1* and *Cdx4* were known direct targets of canonical Wnt signaling (Pilon et al., 2006; Pilon et al., 2007) and Cdx proteins were also known to physically and functionally interact with Lef/Tcf-ß-Catenin (i.e. the nuclear effector complex of the canonical Wnt pathway) to activate transcription via Lef/Tcf response elements (Beland et al., 2004); 3) The presence/binding of the Lef/Tcf-ß-Catenin complex onto any *Pax3* CRM was not reported; and 4) We had identified several putative CdxBS in the NCE region of the 1.6 kb *Pax3* proximal promoter (Milewski et al., 2004), which was also reported to contain not only tissue-specific activity but also to yield enough *Pax3* expression levels to rescue the NC malformations of $Pax3^{Sp/Sp}$ mutants (Li et al., 1999).

0.7.1 General objective

Demonstrate a role for Cdx proteins in the regulation of trunk NC development in the mouse and identify Cdx targets in the trunk NC-GRN.

0.7.2 Specific objectives

1. Analyze global Cdx functions in NC development *in vivo*.

2. Demonstrate a Cdx role as intermediates in Wnt-mediated induction of *Pax3* expression at the PNP border and the molecular mechanisms implicated in this regulation.

0.7.3 General methodology

At the beginning of this project, there was very limited information in support of a link between *Cdx* genes and NC development in vertebrates. One study reported that *Cdx* loss-of-function in zebrafish results in a reduced number of Rohon-Beard neurons, which share a common precursor with NCC (Skromne et al., 2007). The only other supporting data was the report of spinal ganglia fusions in $Cdx1^{-/-}$::$Cdx2^{+/-}$ double mouse mutants, although it was not clear whether *Cdx* genes impacted on NCC AP-patterning in a tissue-autonomous manner or not (van den Akker et al., 2002). We reasoned that Cdx functions in the NC may have been masked by functional overlap and/or early embryonic lethality. In order to circumvent these limitations, we used a conditional dominant negative approach. Such a dominant negative strategy was successfully used in ascidians to uncover a direct role for the Ci-Cdx ortholog in NT closure (Mita and Fujiwara, 2007). We then engineered a Flag-tagged obligatory repressor fusion protein ($_{FLAG}$-EnRCdx1) consisting of the repressor domain of *Drosophila* Engrailed (EnR) fused to the DNA binding domain (homeodomain) of Cdx1. The Cdx homeodomain is highly conserved (around 90%) between Cdx1, Cdx2 and Cdx4 proteins and is used for DNA target recognition and protein-protein interactions (Beland et al., 2004; Lafontaine et al., 2012; Marom et al., 1997). Therefore, via the Cdx1 homeodomain, EnRCdx1 is able to compete with endogenous Cdx proteins for Cdx binding sites (CdxBS) as well as for interacting partners.We used this tool to perturb Cdx function *in vitro* as well as in the NC *in vivo* via targeting of EnRCdx1 coding sequences to the *ROSA26* locus and generation of a conditional Cre/loxP-dependent EnRCdx1 mouse model.

CHAPTER I

Caudal-related homeobox (Cdx)-dependent integration of canonical Wingless/int1-related (Wnt) signals on a well-conserved neural crest enhancer of the proximal *Paired-box 3* (*Pax3*) promoter

(Published in the Journal of Biological Chemistry, 2012, vol 287, pages 16623-35)

Oraly Sanchez-Ferras, Baptiste Coutaud, Taraneh Djavanbakht Samani, Isabelle Tremblay, Ouliana Souchkova and Nicolas Pilon.

Molecular Genetics of Development, Department of Biological Sciences and BioMed Research Center, Faculty of Sciences, University of Quebec at Montreal (UQAM)

To whom correspondence should be addressed: Nicolas Pilon, Email: pilon.nicolas@uqam.ca

Author contribution:

Oraly Sanchez-Ferras: conception, design, acquisition, analysis and interpretation of data, drafting and revision of the article; (experiments: RT-PCR in Wnt3a-treated P19 cells, whole mount in situ hybridization of *Pax3* expression in embryos, analysis of genetic interaction between *Cdx1* and *Pax3 in vivo*, co-transfection analyses in P19 and Neuro2a cells, electrophoretic mobility shift assays and mouse transgenic studies). *Baptiste Coutaud*: acquisition and analysis of data (experiments: co-transfection analyses in P19 and N2a cells). *Taraneh Djavanbakht Samani*: Acquisition and analysis of data (experiments: RT-PCR and co-transfection analyses in N2a cells). *Isabelle Tremblay*: acquisition and analysis of data (experiments: co-transfection analyses in N2a and P19 cells). *Ouliana Souchkova*: generation of RNA probes for *in situ* hybridization. *Nicolas Pilon:* conception, design, acquisition, supervision of the work and revision of the manuscript; (experiments: chromatin immunoprecipitation analysis in N2a cells; generation and validation of the EnRCdx1 construct, whole-mount *in situ* hybridization and immunohistochemistry analyses, RT-PCR analyses in N2a cells).

1.1 Summary

One of the earliest events in neural crest development takes place at the neural plate border and consists in the induction of *Pax3* expression by posteriorizing Wnt/ß-Catenin signaling. The molecular mechanism of this regulation is not well understood but several observations suggest a role for posteriorizing Cdx transcription factors (Cdx1/2/4) in this process. *Cdx* genes are known as integrators of posteriorizing signals from Wnt, RA and FGF pathways. In this work, we report that Wnt-mediated regulation of murine *Pax3* expression is indirect and involves Cdx proteins as intermediates. We show that *Pax3* transcripts co-localize with Cdx proteins in the posterior neuroectoderm and that neural *Pax3* expression is reduced in *Cdx1*-null embryos. Using Wnt3a-treated P19 cells and neural crest-derived Neuro2a cells, we demonstrate that *Pax3* expression is induced by the Wnt-Cdx pathway. Co-transfection analyses, electrophoretic mobility shift assays, chromatin immunoprecipitation as well as transgenic studies further indicate that Cdx proteins operate via direct binding to an evolutionary conserved neural crest enhancer of the *Pax3* proximal promoter. Taken together, these results suggest a novel neural function for Cdx proteins within the gene regulatory network controlling neural crest development.

1.2 Introduction

Pax3 is a paired-box homeodomain transcription factor essential for normal neural crest (NC), neural tube (NT) and skeletal muscle development. In the mouse, homozygous *Pax3* loss-of-function, as seen in *Splotch* (*Sp*) mutants, leads to early embryonic lethality due to its role in NC cells (NCC). While $Pax3^{+/Sp}$ mice only display pigmentation anomalies (white belly spot), most $Pax3^{Sp/Sp}$ embryos die around embryonic day (e)14.0 due to a severe NC defect leading to lack of outflow tract septation and heart failure (Li et al., 1999). $Pax3^{Sp/Sp}$ also display other severe anomalies including spinal ganglia malformations, intestinal aganglionosis and

posterior NT defects (spina bifida) (Auerbach, 1954). At the molecular level, Pax3 plays a critical role in the gene regulatory network controlling NCC development downstream of canonical Wnt signals. Together with members of the Msx, Dlx and Zic families, Pax3 specifies the neural plate border and promotes induction of NCC (Betancur et al., 2010; Meulemans and Bronner-Fraser, 2004; Monsoro-Burq et al., 2005). At later stages, Pax3 also controls survival of dorsal NT progenitors through stimulation of p53 degradation (Pani et al., 2002; Wang et al., 2011). In humans, heterozygous *PAX3* mutations have been associated with Waardenburg syndrome which is characterized by NC defects such as cranio-facial and pigmentary anomalies (Baldwin et al., 1992; Baldwin et al., 1995; Tassabehji et al., 1992).

Neural *Pax3* expression begins at the early somite stage (around e 8.25) prior to initiation of NT closure. At this stage, *Pax3* transcripts are detected on the lateral borders of both the anterior and posterior neural plate (ANP and PNP, respectively) (Goulding et al., 1991). Following neural plate bending in a closed NT (from e8.5 onwards), *Pax3* transcripts are then detected in the dorsal neuroectoderm (including pre-migratory NCC) in an almost continuous manner along the anterior-posterior (A-P) axis. Indeed, strong *Pax3* expression is detected in two large domains extending 1) from the forebrain down to rhombomere 4 and 2) from rhombomere 6 down to the rostral half of the PNP. Although detectable, *Pax3* expression is clearly much weaker in rhombomere 5. At later stages, caudal *Pax3* expression follows progression of posterior elongation and is maintained in the dorsal half of the closed neural tube until e14.5. *Pax3* expression is also detected in a subset of migratory NCC contributing to the cardiac outflow tract, peripheral and enteric nervous systems, as well as melanocytes but is generally downregulated as NCCs differentiate.

Regulatory sequences sufficient to mediate induction and dorsal restriction of *Pax3* expression in the hindbrain and trunk are contained in the proximal 1.6 kb promoter (Natoli et al., 1997). Deletion analysis of this promoter has revealed that a block of

674bp containing two evolutionary conserved regions of approximately 250bp, called Neural Crest Enhancer (NCE) 1 and 2, is sufficient to drive expression in the dorsal NT and NCC (Milewski et al., 2004). NCE1 bears Pbx binding sites (activated by Pbx1-containing transcriptional complexes) which appear specifically required to control *Pax3* expression in the hindbrain (Chang et al., 2008; Pruitt et al., 2004). NCE2 contains a Tead binding site that was shown to be required for the activity of the whole 674bp NCE in e10.5 transgenic embryos (Milewski et al., 2004). In addition, both NCE1 and NCE2 contain a binding site for Pou class III members and mutation of these sites leads to reduced NT activity of the 1.6kb promoter in e9.5 transgenic embryos (Pruitt et al., 2004). On the other hand, *Pax3* expression is induced by posteriorizing Wnt signals and dorsally restricted in response to dorso-ventral patterning signals such as Sonic Hedgehog (Shh) (Bang et al., 1999; Bang et al., 1997; de Croze et al., 2011; Goulding et al., 1993; Monsoro-Burq et al., 2005; Taneyhill and Bronner-Fraser, 2005). However, how the canonical Wnt and Shh signals are integrated at the transcriptional level is still unclear.

The vertebrate *Cdx* genes (*Cdx1*, *Cdx2* and *Cdx4*) are related to *Drosophila caudal* (*cad*) (Brooke et al., 1998) and their gene products have conserved the ancestral ability to specify the posterior embryo and pattern the A-P axis. Murine *Cdx* genes are sequentially activated in ectodermal and mesodermal cells of the primitive streak around e7.25, with *Cdx1* activated first and *Cdx4* activated last (Beck et al., 1995; Gamer and Wright, 1993; Meyer and Gruss, 1993). At e8.5, all *Cdx* genes are expressed in the caudal embryo and form a nested set along the A-P axis. While *Cdx1* and *Cdx2* exhibit an almost perfect overlap in expression around the hindbrain/spinal cord boundary, *Cdx4* is expressed slightly more posterior at this stage. The most anterior domain of *Cdx* expression appears to be restricted to the dorsal NT, with Cdx1 protein expressed in NCC emigrating from this domain (Gaunt et al., 2005; Meyer and Gruss, 1993). This anterior limit of expression regresses concomitantly with axial elongation, persisting in the caudal embryo until e10.5, for *Cdx1* and *Cdx4,*

and e12.5 for *Cdx2*. All three *Cdx* genes are also expressed in the developing hindgut epithelium with *Cdx1* and *Cdx2* expression maintained postnatally (Beck, 2002).

Cdx gene expression is regulated by posteriorizing signals from Wnt, RA and FGF pathways in multiple species (Bel-Vialar et al., 2002; Faas and Isaacs, 2009; Houle et al., 2000; Houle et al., 2003; Ikeya and Takada, 2001; Keenan et al., 2006; Lengerke et al., 2008; Lickert et al., 2000; Pilon et al., 2006; Pilon et al., 2007; Prinos et al., 2001; Shimizu et al., 2005). Among these posteriorizing signals, the evolutionary conserved role of the canonical Wnt pathway on *Cdx* regulation is the best characterized. Indeed, both *Cdx1* and *Cdx4* have been clearly identified as direct targets of the Wnt-ß-Catenin pathway (Lickert et al., 2000; Pilon et al., 2006; Prinos et al., 2001). Moreover, other data suggest that *Cdx2* is also responsive to canonical Wnt signals, although evidence for a direct regulation is sparse (Benahmed et al., 2008; He et al., 2008; Joo et al., 2010; Marikawa et al., 2009; Nordstrom et al., 2006; Saegusa et al., 2007; Zhao and Duester, 2009). In addition, Cdx proteins can interact with the Lef1-ßCatenin transcriptional effector of the canonical Wnt pathway (Beland et al., 2004).

Our understanding of Cdx function has long been hampered by the functional overlap between Cdx members and the vital role of *Cdx2* during implantation (Savory et al., 2009b; Strumpf et al., 2005). Recent development of a conditional *Cdx2* allele (Gao et al., 2009; Savory et al., 2009a) and analysis of *Cdx* double mutants (Savory et al., 2011a; van den Akker et al., 2002; van Nes et al., 2006; Young et al., 2009) has demonstrated important regulatory roles for Cdx proteins in several processes during mouse embryogenesis. In the mesoderm, Cdx proteins regulate axial patterning, axial elongation and somitogenesis in addition to placentogenesis and hematopoiesis. An important part of the Cdx function is executed through direct regulation of diverse *Hox* genes (Pilon et al., 2007; Subramanian et al., 1995; van den Akker et al., 2002; van Nes et al., 2006; Wang et al., 2008b; Young et al., 2009). However, recent work

has shown that a significant proportion of the Cdx function is also fulfilled via direct regulation of several non-*Hox* targets (Grainger et al., 2011; Savory et al., 2009a). In the endoderm, Cdx proteins are involved in intestinal patterning and cell differentiation via *Hox*-dependent and -independent mechanisms (Crissey et al., 2011; Gao et al., 2009; Grainger et al., 2010; Suh et al., 1994; Verzi et al., 2011).

The function of Cdx members in the neuroectoderm is less well understood. As in the mesoderm and endoderm, Cdx proteins control neuroectoderm A-P patterning via *Hox*-dependent and -independent mechanisms (Bel-Vialar et al., 2002; Isaacs et al., 1998; Nordstrom et al., 2006; Skromne et al., 2007; Sturgeon et al., 2011). Multiple gain- or loss-of-function mouse models have also demonstrated that Cdx proteins are important for proper formation of the NT and NC-derived spinal ganglia but whether they do so in a tissue-autonomous or -non-autonomous (via the mesoderm) manner remains an open question (Charite et al., 1998; Gaunt et al., 2008; Savory et al., 2011a; van de Ven et al., 2011; van den Akker et al., 2002). Nevertheless, analysis of neural specific *Cdx* loss-of-function in ascidian embryos demonstrates that Cdx proteins may impact on NT and NC formation in a tissue-autonomous manner. Indeed, such mutant embryos display NT defects as well as absence of pigment cells, which are derived from NC-like cells (Jeffery et al., 2004; Mita and Fujiwara, 2007).

In this study, we investigated the possibility that Cdx proteins might convey the posteriorizing Wnt signals to *Pax3*. We found that murine *Pax3* is in fact an indirect Wnt target and that Cdx proteins can directly activate neural *Pax3* expression at least via the previously described NCE2 (Milewski et al., 2004). Altogether, our data strengthen the idea that Cdx members are involved in both NT and NCC development in a tissue-autonomous manner downstream of Wnt signals.

1.3 Experimental procedures

1.3.1 Ethics Statement

Experiments involving mice were performed following Canadian Council of Animal Care (CCAC) guidelines for the care and manipulation of animals used in medical research. Protocols involving the manipulation of animals were approved by the institutional ethics committee of the University of Quebec at Montreal (comité institutionnel de protection des animaux (CIPA)); Reference number 0511-R2-648-0512).

1.3.2 Generation and analysis of mice

Cdx1$^{-/-}$ mutant mice (Subramanian et al., 1995) were kindly provided by D. Lohnes (University of Ottawa, Canada). *Pax3*$^{+/Sp}$ mice were obtained from the Jackson Laboratories (Bar Harbor, Maine). *Cdx1*$^{+/-}$*Pax3*$^{+/Sp}$ mutants were generated by *Pax3*$^{+/Sp}$ and *Cdx1*$^{-/-}$ intercrosses. *Cdx1*$^{-/-}$*Pax3*$^{+/Sp}$ mutants were then generated by crossing *Cdx1*$^{+/-}$*Pax3*$^{+/Sp}$ and *Cdx1*$^{-/-}$ mice.

Transgenes carrying the *lacZ* gene under control of the *Hsp68* minimal promoter (Kothary et al., 1989) and either a wild-type or Cdx binding site (CdxBS)-mutated NCE2 were prepared and transgenic embryos generated from injected B6C3 oocytes according to standard techniques (Nagy et al., 2003). In order to facilitate identification of transgenic embryos and provide a positive control for transgene expression, a previously described *Gata4p*-GFP transgene (Pilon et al., 2008) was co-injected with each lacZ construct. Nine days after microinjection, foster mothers were sacrificed and embryos were collected and individually analyzed for GFP and ß-Galalactosidase activity. For X-Gal staining, embryos were fixed in 2% PFA for 15 min., washed twice with PBS and incubated overnight in staining solution ($K_3Fe(CN)_6$ 5mM, $K_4Fe(CN)_6$ 5mM, $MgCl_2$ 2mM, NP-40 0.01%, sodium

deoxycholate 0.01% in PBS) containing 1mg/ml of X-Gal. Staining reactions were carried out overnight at 37 °C.

Whole-mount *in situ* hybridization and immunohistochemistry were performed as previously described (Pilon et al., 2006; Savory et al., 2009b). Mice were mated overnight and noon of the day of vaginal plug detection was considered as e0.5. Embryos to be compared were stage-matched according to established criteria (Kaufman, 1992) and processed in parallel. The probe for *in situ* hybridization of *Pax3* mRNA was generated from a previously described plasmid (Goulding et al., 1991), kindly provided by J. Epstein. Immunohistochemistry was performed using polyclonal antisera to Cdx1 and Cdx2 (Savory et al., 2009b).

All embryo images were taken with a Leica DFC 495 camera mounted on a Leica M205 FA stereomicroscope (Leica Microsystems Canada).

1.3.3 Chromatin immunoprecipitation (ChIP) analysis

ChIP assays in Neuro2a (N2a) cells were performed as previously described (Pilon et al., 2006) using anti-FlagM2 antibody (Sigma). PCR amplifications were performed with GoTaq DNA polymerase (Promega) and consisted of 30 cycles of 30s at 96°C, 30s at 60°C and 30s at 72°C. The primers used for this study were: *Pax3* NCE2 Forward (5'-GGCACAATGGTACCTTCTCTAAGG) and Reverse (5'-AAGCTTCCCTTCTGAGAAGCGGGGACTTTAAA); *Pax3* exon 7 Forward (5'-CCGTGTCAGATCCCAGTAGCAC) and Reverse (5'-CTGAGGTGAAAGGCCATTGCCG). PCR products were resolved on a 1.5% agarose gel.

1.3.4 Electrophoretic mobility shift assays (EMSA)

Pax3 NCE2 was scanned by using 8 double-stranded end-labeled oligonucleotides and tested for Cdx binding by EMSA. Five micrograms of nuclear extracts from

mock or Flag-Cdx1 transfected Cos7 cells were used in each reaction as previously described (Pilon et al., 2006). Supershifts were performed using 1 μg of anti-FlagM2 antibody (Sigma). Specificity of binding was assessed by competition with a 100-fold excess of unlabeled WT or mutated cold probes. The oligonucleotides comprising mutated CdxBS1 and CdxBS3 sequences were as described for site-directed mutagenesis. The upper strands of each wild-type or mutated double-stranded probe as well as of a probe harbouring a consensus Cdx binding site (as positive control) are summarized in Table S1.

1.3.5 Plasmid constructs

For generation of the Engrailed-Cdx1 (EnRCdx1) fusion construct, the coding sequence of the *Drosophila* Engrailed repressor domain (Jaynes and O'Farrell, 1991) was subcloned upstream of a PCR product corresponding to the DNA binding domain (homeodomain) of Cdx1 ended by a Stop codon (See Figure 1.8). Flag-tagged Cdx1 expression vector has been described previously (Beland et al., 2004). Flag-tagged EnRCdx1 and EnR constructs were generated by subcloning the relevant sequences into a modified *pCEP4* plasmid (Invitrogen; (Beland et al., 2004)). Cdx1, Cdx2, Cdx4, EnR and EnRCdx1 bicistronic expression vectors were generated by subcloning the respective cDNA into the pIRES2-EGFP vector (Clontech).

Pax3 proximal promoter sequences were obtained by PCR amplification in accordance to previous work (Li et al., 1999; Milewski et al., 2004). Oligonucleotide sequences are available upon request. These PCR products were cloned in pGEM-T vector and validated by sequencing. Luciferase reporter constructs bearing the *Pax3* minimal promoter with or without various lengths of the 1.6 kb 5' upstream sequences were generated by subcloning these sequences into pXP2 (Nordeen, 1988).

Point mutations were introduced into each of the three CdxBS of NCE2 by using the Quick Change Multisite-directed mutagenesis kit (Stratagene), according to

manufacturer's instructions. Sequences of the oligonucleotides used for site-directed mutagenesis were CdxBS1 5'-CAGCAGTTTAGTCTGAATGCCATAATAccTTCCTGAGAAC; CdxBS2 5'-CTAGCCAAGACGTTGCTTCTTcgATTTTTCCAGCAGTTTA; CdxBS3 5'-AAGGACAGACAGTCTcgACAACACTCCTGGCGTCATATCC (point mutations are denoted in lower-case letters).

1.3.6 Cell culture and transfection analysis

P19 cells were propagated in AMEM supplemented with 7.5% FBS and 2.5% NCS, N2a cells in EMEM supplemented with 10% FBS and Cos7 cells in DMEM supplemented with 10% FBS. All transfections were performed using GeneJuice reagents (Novagen) in accordance with manufacturer's instructions.

For time course analysis, P19 cells were seeded in 6 well plates ($2x10^5$ cells per well) and the following day standard culture medium was replaced by a Wnt3a- or a Ctl-conditioned medium (Pilon et al., 2006). Cells were harvested before treatment (t=0h) and 2h, 4h, 8h, 12h or 18h post-treatment, snap frozen and stored at -80°C prior to RT-PCR analysis. To assess the requirement for *de novo* protein synthesis, P19 cells were pre-treated for 30 min with 30 µg/mL of cycloheximide (CHX) or with the vehicle (DMSO) alone and then treated for 24h with Wnt3a or Ctl-conditioned medium in the presence of 1 µg/mL of CHX. Afterwards, cells were harvested, snap frozen, and stored as described above prior to RT-PCR analysis. To inhibit Cdx function, P19 cells were seeded in 100-mm plates ($2x10^6$ cells per plate) and transiently transfected with 6 µg of EnRCdx1-IRES-GFP or the negative control EnR-IRES-GFP expression vector. Thirty hours after transfection, cells were treated with Wnt3a- or Ctl-conditioned medium and cultured for another twenty hours. GFP-positive cells were then recovered by FACS and analyzed by RT-PCR.

To modulate Cdx activity in N2a cells, transient transfections with 6 µg of EnRCdx1-IRES-GFP, Cdx1-IRES-GFP, Cdx2-IRES-GFP or Cdx4-IRES-GFP expression vectors were performed in 100-mm tissue culture plates ($2x10^6$ cells per plate). Forty-eight hours after transfection, GFP-positive cells were recovered by FACS and analyzed by RT-PCR.

For luciferase reporter assays, N2a ($8x10^4$ cells per well) and/or P19 ($3x10^4$ cells per well) cells were transfected in 24-well plates with 100 ng of *Pax3*p-Luciferase reporter construct alone or with increasing amounts (1.25 to 5 ng) of Cdx expression vectors; or with the maximum amount of Cdx expression vector and increasing amounts (2.5 to 20 ng) of EnRCdx1 expression vector. When required, an empty expression vector was also included to insure a total of 125 ng of DNA per well. For analysis of deleted or mutated *Pax3*p-Luciferase reporter constructs, a fixed concentration (5 ng) of Cdx2 expression vector was used. All transfections were performed at least three times in triplicate. Forty-eight hours after transfection, cells were disrupted in 100 µl of lysis buffer (0.1 M Tris [pH 8.0], 1% Igepal, 1 mM dithiothreitol) and assessed for luciferase activity with a Berthold LB9507 luminometer (Berthold Technologies).

1.3.7 RNA extraction and RT-PCR analysis

Total RNA was isolated from frozen cell pellets by using the RNeasy kit (Qiagen) and cDNA was synthesized using a polydT oligonucleotide and Superscript II Reverse-Transcriptase (Invitrogen) in accordance with manufacturer's protocols. PCR amplifications were then performed with Platinum Taq DNA polymerase (Invitrogen) and consisted of 25 to 35 cycles of 30 sec at 96°C, 30 sec at 60°C, and 45 sec at 68°C. Amplified bands were size fractioned on a 2% agarose gel. The primers used in this study were: Cdx1 (forward: 5'-GCAAGTCCGAGCTGGCTGCTA, reverse: 5'-GGGTAGAAACTCCTCCTTGACG); Cdx2 (forward: 5'-

CCACACTTGGGCTCTCCGAGA, reverse: 5'-GGGTCACTGGGTGACAGTGGA); Cdx4 (forward: 5'-AGTTAACCTGGGCCTTTCTGA, reverse: 5'-ATTCAGAAACTATGACCTGCTGTATC), Pax3 (forward: 5'-CCTGCCAACATACCAGCTGTCG, reverse: 5'-CTGAGGTGAAAGGCCATTGCCG); Gapdh (forward: 5'-TCCTGCACCACCAACTGCTTAGC, reverse: 5'-AGGTCCACCACCCTGTTGCTGTA). All oligonucleotides were designed to encompass an intron allowing the detection of contaminating genomic DNA by the presence of a larger band. As an additional control, PCR was also performed with RNA that had not been reverse-transcribed. Statistical analysis was carried out using GraphPad Prism software version 5.0. For the comparison of groups, a paired Student *t*-test was used (two-tailed *p* value, α=0.05). Differences between means were classed as not significant (n.s., $p>0.05$) or significant (*, $p<0.05$).

1.4 Results

1.4.1 Cdx members and Pax3 are co-expressed in the caudal neuroectoderm

To validate our hypothesis that Cdx proteins are good candidates to induce *Pax3* expression in the caudal neuroectoderm, we compared the expression pattern of Cdx proteins and *Pax3* mRNA. This analysis revealed extensive Cdx-*Pax3* overlap in the caudal neuroectoderm of e8.5 and e9.5 embryos (Figure 1.1), in accordance with the well documented Cdx and *Pax3* expression patterns (Beck et al., 1995; Goulding et al., 1991; Meyer and Gruss, 1993). In e8.5 embryos, *Pax3* transcripts are detected at both the PNP border and the site of initiation of NT closure (around the level of the fifth somite) where Cdx proteins are strongly expressed (Figure 1.1C,D). In e9.5 embryos, *Pax3* and Cdx overlapping expression domains are displaced caudally in parallel with progression of axial elongation, NCC induction and NT closure (Figure

1.1E-H). In brief, Cdx spatio-temporal expression patterns are consistent with a role in the induction of *Pax3* expression in the caudal neuroectoderm.

1.4.2 Pax3 neural expression is regulated by Cdx proteins

To determine whether *Pax3* is a Cdx target gene, we evaluated its expression in e8.5 *Cdx1*-null embryos. This analysis revealed that *Pax3* expression is slightly reduced in the dorsal NT of 4- to 6-somite stage $Cdx1^{-/-}$ embryos (Figure 1.2A). Our data also suggest that this effect is transient as no difference in *Pax3* expression is noted at later stages (>8 somites) (data not shown). To evaluate if this slight and transient reduction in *Pax3* expression is functionally significant, we generated and analyzed an allelic series of *Cdx1-Pax3* compound mutants. $Cdx1^{+/-}Pax3^{+/Sp}$ and $Cdx1^{-/-}Pax3^{+/Sp}$ animals were obtained at expected mendelian ratios and were similar to $Pax3^{+/Sp}$ animals, exhibiting a white belly spot (data not shown). Given the known functional overlap between Cdx members (Faas and Isaacs, 2009; Savory et al., 2009b; van den Akker et al., 2002; van Nes et al., 2006), this raises the possibility that the presence of Cdx2 and Cdx4 ensures that *Pax3* expression levels do not significantly fall below 50% in these compound mutants. Therefore, the transient reduction in *Pax3* mRNA levels observed in the dorsal NT of $Cdx1^{-/-}$ embryos most likely reflects the fact that *Cdx1* is expressed slightly earlier and more anteriorly than other *Cdx* members. In this regard, it is noteworthy that a similar transient effect has been recently reported for expression of the Cdx neural target *Mafb* in *Cdx1*-null embryos (Sturgeon et al., 2011). Taken together, this suggests that all Cdx proteins might be involved in the control of *Pax3* expression.

To circumvent the functional overlap of Cdx members, we generated a Flag-tagged dominant negative Cdx protein consisting of the repressor domain of *Drosophila* Engrailed fused to the DNA binding domain (homeodomain) of Cdx1 (EnRCdx1; see Figure 1.8). Given that the homeodomain is highly conserved (more than 90%) among Cdx members, EnRCdx1 is expected to recognize every Cdx target genes and

inhibit their expression. We used the EnRCdx1 dominant negative tool to assess whether endogenous *Pax3* expression is affected by modulation of Cdx activity in N2a cells. NC-derived N2a cells are a good model to study the regulation of the endogenous *Pax3* promoter by Cdx, as they co-express all three *Cdx* members as well as *Pax3* (Figure 1.2B). N2a cells were transiently transfected with expression constructs encoding Cdx1, Cdx2, Cdx4 or the dominant negative EnRCdx1 and endogenous *Pax3* expression was evaluated by semi-quantitative RT-PCR. As shown in Figure 1. 2C, *Pax3* expression levels are significantly altered by modulation of Cdx activity in N2a cells. Over-expression of any of the three Cdx members results in a robust increase whereas over-expression of EnRCdx1 leads to a complete knockdown of *Pax3* expression. Therefore, these data indicate that *Pax3* is a Cdx target gene in N2a cells.

1.4.3 Pax3 is induced by the Wnt-Cdx pathway in undifferentiated P19 cells

Previous studies have reported that neural *Pax3* expression is induced by canonical Wnt signals (Bang et al., 1999; de Croze et al., 2011; Monsoro-Burq et al., 2005; Taneyhill and Bronner-Fraser, 2005). To evaluate the potential contribution of Cdx proteins in this process, we performed a series of RT-PCR analyses in P19 embryocarcinoma cells cultured in presence or absence of Wnt3a-conditioned medium (W3a-CM), which is known for activating the canonical pathway (Pilon et al., 2006; Shibamoto et al., 1998). The N2a cell line could not be used for this assay as the canonical Wnt pathway is constitutively activated in these cells owing to Wnt7a autocrine regulation (Colombres et al., 2008; Shi et al., 2010). On the other hand, P19 cells have been used previously to study the regulation of *Pax3* expression in the context of neural differentiation (Milewski et al., 2004; Natoli et al., 1997; Pruitt et al., 2004) and activation of the Wnt/ß-Catenin pathway (Petropoulos and Skerjanc, 2002). As shown in Figure 1.3A, a time-course analysis first revealed that expression of each *Cdx* member is rapidly induced (2-4h) following treatment with W3a-CM. Under the same conditions, *Pax3* expression is also induced but

considerably delayed (18h). This delay being suggestive of an indirect regulation, we then directly assessed the necessity for an intermediary factor by inhibiting *de novo* protein synthesis with the protein synthesis inhibitor cycloheximide (CHX). *Cdx1* and *Cdx4* are known direct targets of Wnt/ß-Catenin signaling (Lickert et al., 2000; Pilon et al., 2006; Pilon et al., 2007; Prinos et al., 2001). Accordingly, *Cdx1* and *Cdx4* induction by W3a-CM is not affected by inhibition of protein synthesis (Fig. 3B). Interestingly, *Cdx2* induction was also found to be independent of *de novo* protein synthesis, suggesting that *Cdx2* is a direct target of Wnt/ß-Catenin signaling in this model. In marked contrast, *Pax3* induction is blocked by treatment with CHX, demonstrating that a protein intermediary is needed to convey canonical Wnt signals to the *Pax3* promoter (Figure 1.3B). Of note, such outcome is also supported by transient co-transfection assays indicating that Lef1/ß-Catenin complexes are very weak activators of the proximal 1.6kb promoter of *Pax3* in N2a or P19 cells (1.6-fold; data not shown).

To determine whether Cdx members act as intermediaries between canonical Wnt signals and *Pax3*, we knocked-down Cdx function in P19 cells using the EnRCdx1 dominant negative protein. This analysis indicated that overexpression of EnRCdx1, but not that of EnR alone, strongly impairs induction of *Pax3* by Wnt3a (Figure 1.3C), suggesting that Cdx function is required for this regulation. Taken together, these results indicate that the neural plate border specifier *Pax3* is an indirect target of Wnt/ß-Catenin signaling and implicate Cdx members as mediators of Wnt inductive signals.

1.4.4 Identification of Cdx binding sites in the proximal Pax3 promoter

Previous studies have demonstrated that the proximal 1.6kb promoter of *Pax3* recapitulates endogenous posterior expression of *Pax3* in the dorsal NT and NCC (Milewski et al., 2004; Natoli et al., 1997). In order to understand the mechanism of Cdx-mediated regulation of *Pax3* expression, we performed a series of co-

transfection assays in N2a cells. To assess whether Cdx proteins transactivate the proximal 1.6 kb *Pax3* promoter, we generated a luciferase reporter construct driven by these regulatory sequences (*Pax3*p1.6kb-Luc) and evaluated the effect of Cdx1, Cdx2, Cdx4 or EnRCdx1 expression on its activity. As shown in Figure 1.4A-C, all three Cdx proteins robustly induce this reporter in a dose-dependent manner, and each induction is strongly repressed by EnRCdx1. It is also interesting to note that, although each Cdx is equally expressed (data not shown), Cdx2 elicits the strongest response (32-fold) followed by Cdx1 (11-fold) and Cdx4 (5-fold). Such observation is in accordance with previous work and suggests that Cdx members do not exhibit the same strength of transactivation (Savory et al., 2011b).

The Cdx-dependent transcriptional response of the proximal 1.6 kb promoter correlates with the presence of five putative Cdx binding sites identified via bioinformatic (MatInspector, Genomatix) and manual analyses. Interestingly, four of these potential Cdx binding sites are located in the previously described NCE1 and NCE2. To better define the sequence elements that mediate Cdx transactivation, a promoter deletion analysis was carried out in N2a cells and Cdx2 was used to assay Cdx responsiveness. This assay revealed that the NCE2 region, which contains three putative Cdx binding sites, mediate most of Cdx2 transcativation (Figure 1.4D). Accordingly, ChIP-PCR assays performed in N2a cells indicate that both Flag-tagged Cdx1 and EnRCdx1 proteins are present on the endogenous *Pax3* NCE2 (Figure 1.4E).

EMSA was then used to verify whether putative Cdx binding sites (called CdxBS1, CdxBS2 and CdxBS3) contained in NCE2 can be directly bound by a Cdx protein. In order to rule out the possibility that other unpredicted Cdx binding sites might also exist, we first scanned the whole NCE2 sequences with 8 overlapping double-stranded oligonucleotide probes (Figure 1.9 and Table 1.1). Incubation of these probes with nuclear extracts from Cos7 cells overexpressing Flag-Cdx1 revealed that

Cdx1 preferentially binds to probes bearing either CdxBS1 or CdxBS3. Cdx1 binding to the probe bearing CdxBS2 was very weak while no binding was observed for the remaining probes. The presence of Flag-Cdx1 in the complex formed with probes containing CdxBS1 or CdxBS3 was demonstrated by supershift with an anti-Flag antibody (Figure 1. 5B). Moreover, specificity of Cdx binding to CdxBS1 and CdxBS3 was confirmed by the absence of competition with cold probes bearing point mutations in the predicted elements (Figure 1.5B). Taken altogether, our data demonstrate that Cdx proteins directly regulate *Pax3* expression at least via functional binding sites contained in *Pax3* NCE2.

1.4.5 Cdx binding sites are essential for Pax3 NCE2 activity

To determine whether Cdx binding sites identified by EMSA are functionally important for *Pax3* NCE2 activity, we generated a series of NCE2-luciferase reporter constructs containing the mutated elements (alone or in combination) and evaluated their responsiveness to Cdx2 in co-transfection assays. In P19 cells, single and double mutations of CdxBS reduce Cdx2-mediated transactivation of the NCE2 reporter while simultaneous mutation of all three CdxBS results in complete loss of Cdx2-mediated transactivation (Figure 1.5C). In accordance with our EMSA data, this analysis also revealed that the contribution of CdxBS1 and CdxBS3 to NCE2 activity is more important than that of CdxBS2. Similar results were obtained in N2a cells, although it can be noted that mutation of CdxBS1 does not considerably inhibit the Cdx2 responsiveness of *Pax3* NCE2 in these cells. Interestingly, it can also be noted that the order of transactivation of the wild-type NCE2 reporter is higher in N2a (approximately 14-fold) than in P19 cells (approximately 6-fold) and that the triple-mutant construct remains slightly more active in N2a than P19 cells. These observations suggest that Cdx proteins, in addition to directly binding the *Pax3* NCE2, might also be recruited to and/or stabilized on this enhancer via an interaction with a neural specific factor also required for strong activation.

To verify the importance of all three Cdx binding sites of *Pax3* NCE2 *in vivo*, we generated *lacZ* reporter constructs driven by either wild-type or mutated *Pax3* NCE2 upstream of the *Hsp68* minimal promoter and evaluated ß-Galactosidase activity in transient e9.5 transgenic embryos (Figure 1.6). This analysis first revealed that this short enhancer of 245 bp is sufficient to recapitulate *Pax3* expression in the dorsal NT and pre-migratory NCC of the caudal embryo (Figure 1.6A; compare with Figure 1.1H). Importantly, we also found that point mutations in all three Cdx binding sites abolish NT and NCC expression of the NCE2 reporter (Figure 1.6B). Of note, this lack of expression of the mutated NCE2 reporter is not due to transgene integration in a repressive chromatin region since expression of a co-injected *Gata4p-GFP* transgene is not affected (Figure 1.6C-D). Therefore, these results demonstrate that intact Cdx binding sites in *Pax3* NCE2 are required to recapitulate caudal and dorsal neuroectoderm-specific expression of *Pax3* in transgenic mice.

1.5 Discussion

We presented data indicating that the dorsal NT/NCC marker *Pax3* is a direct target of the Cdx proteins downstream of canonical Wnt signals. Cdx proteins convey canonical Wnt signals to the proximal *Pax3* promoter through direct binding to Cdx binding sites located in the evolutionary conserved NCE2. These Cdx binding sites are essential for both Cdx-mediated transactivation of NCE2 in cell culture experiments and for expression of a NCE2 reporter in the dorsal NT and NCC of e9.5 transgenic embryos, supporting the existence of the Wnt-Cdx-Pax3 pathway *in vivo*.

1.5.1 Wnt-mediated induction of Pax3 expression at the neural plate border

Although several studies have reported that *Pax3* is a posterior Wnt-induced gene, very little is known regarding the possible mechanism of this regulation (Bang et al., 1999; de Croze et al., 2011; Monsoro-Burq et al., 2005; Taneyhill and Bronner-Fraser, 2005). Our data now indicate that induction of murine *Pax3* expression by canonical Wnt signals is indirect, involving Cdx proteins as intermediaries. In this

regard, it is interesting to note that all three *Cdx* genes are direct targets of Wnt/ß-Catenin signaling in P19 cells. Although this was expected for *Cdx1* and *Cdx4* (Lickert et al., 2000; Pilon et al., 2006; Pilon et al., 2007; Prinos et al., 2001), such outcome was somehow surprising for *Cdx2* given a previous report indicating that exogenous Wnt3a can specifically induce *Cdx1* but not *Cdx2* in embryo culture (Prinos et al., 2001). However, this result is in agreement with more recent work suggesting that *Cdx2* is also a direct target of canonical Wnt signals (Benahmed et al., 2008; He et al., 2008; Joo et al., 2010; Marikawa et al., 2009; Saegusa et al., 2007).

Our data challenge a recent report showing that *Pax3* expression in the dorsal NT can also be regulated by another evolutionary conserved enhancer located in intron 4 (called ECR2) and described as containing multiple putative Lef/Tcf binding sites (Degenhardt et al., 2010). Indeed, it was reported that mutation of these putative binding sites abrogates ECR2 activity in transgenic zebrafishes. However, this mutated transgene was not assayed in the mouse and these putative binding sites were not shown to be bound by Lef/Tcf proteins. Since the consensus Lef/Tcf binding site exhibits rather low binding specificity for HMG-box proteins, this raises the possibility that putative Lef/Tcf binding sites identified in *Pax3* ECR2 are not bound by Lef/Tcf proteins but rather by other HMG-box proteins such as Sox members (Huang et al., 2010; Kormish et al., 2010; Kuwabara et al., 2009; Liu et al., 2010). Alternatively, it is also possible that *Pax3* is a direct Wnt target in the zebrafish and not in the mouse. Regardless of the mechanism operating in other organisms, our CHX experiments indicate that murine *Pax3* is an indirect Wnt target. Moreover, we have found that a luciferase reporter construct driven by ECR2 is, like NCE2, very poorly activated by Lef1/ß-Catenin complexes in transient transfection assays using P19 and N2a cells (data not shown).

Recent work in *Xenopus* embryos also suggests that Wnt-mediated induction of *Pax3* expression at the neural plate border might be controlled by species-specific

mechanisms. Indeed, de Crozé et al. (2011) have reported that *Pax3* expression at the neural plate border is regulated by canonical Wnt signals via both direct and indirect means. This work showed that, although *Pax3* can be directly induced by canonical Wnt signals, the transcription factor AP2a is required as an intermediate for full activation (de Croze et al., 2011). The existence of such a mechanism in the mouse is very unlikely since knockout of all AP2 isoforms (via deletion of exon 5) has been shown to result in cranial neural crest defects that do not involve reduced *Pax3* expression (Schorle et al., 1996).

On the other hand, other studies in *Xenopus* embryos have revealed that the homeobox gene *Gbx2* is a direct downstream target of Wnt/β-catenin signaling acting upstream of *Pax3* at the neural plate border (Li et al., 2009). Similarly to *Cdx* genes, *Gbx2* is a known posteriorizing gene. However, in marked contrast to Cdx proteins, Gbx2 cannot directly activate *Pax3* expression since it acts as a repressor (Heimbucher et al., 2007; Li et al., 2009). Therefore, in this case, it appears that Gbx2 is required to repress an unknown repressor of *Pax3* at the neural plate border. More work will be required to determine whether this mechanism is conserved in the mouse.

1.5.2 Regulation of Pax3 expression via NCE2

We have shown for the first time that the *Pax3* NCE2 alone is sufficient to recapitulate induction as well as dorsal restriction of *Pax3* expression in the caudal NT. Taken together with previous work, this suggests that NCE2 is involved in the posterior whereas NCE1is rather involved in the anterior expression of *Pax3* (Chang et al., 2008; Milewski et al., 2004; Pruitt et al., 2004). As summarized in Fig.7, our *in vitro* and *in vivo* data further indicate that the activity of NCE2 is regulated by the posteriorizing Wnt-Cdx pathway. Given the broad distribution of Cdx proteins in the posterior neuroectoderm, it is currently unclear how NCE2 exhibits dorsally restricted activity. As previously described for *Pax3* expression, this could first be ensured by a

repressive mechanism involving Shh signals emerging from the node, notochord and floorplate (Goulding et al., 1993); implying that Shh-responsive regulatory sequences are contained within NCE2. Such restricted activity of NCE2 might also be due to an interaction between Cdx proteins and a neural plate border co-factor. In this regard, our transfection assays in N2a cells have suggested that Cdx-mediated transactivation of NCE2 might rely in part on the presence of a neural factor. The identity of such co-factor is currently unknow and it is most likely not a transcription factor previously reported to act on NCE2 (Tead2 and Brn1/2) (Milewski et al., 2004; Pruitt et al., 2004). Indeed, although both Tead2 and Brn1/2 have been implicated in the regulation of *Pax3* expression, their expression pattern is not consistent with a role in the induction of *Pax3* expression *in vivo*. On one hand, Tead2 and its co-factor YAP65 are almost ubiquitously expressed at e8.5-e9.5, becoming restricted to neural tissues only after e10.0 (Sawada et al., 2005; Yasunami et al., 1995). On the other hand, the pro-neural factors Brn1 and Brn2 -as well as other Pou classIII members Brn4 and Tst1/Oct6- are not expressed in the PNP (Bouchard et al., 2005; He et al., 1989; Heydemann et al., 2001; Mathis et al., 1992; Monuki et al., 1990; Sugitani et al., 2002). Thus, these observations strongly suggest that Tead2 and Brn1/2 are involved in the maintenance rather than induction of *Pax3* expression. More work will obviously be required to better understand the regulatory mechanisms involved in the dorsal restriction of *Pax3* expression and our data indicate that at least some of them are operating via NCE2.

1.5.3 Novel function for Cdx proteins in caudal neuroectoderm development

Strong *Cdx* expression in the caudal neuroectoderm is highly conserved across chordates. However, Cdx function in this lineage is poorly understood because of functional overlap. Until recently, Cdx proteins were mostly known for their evolutionary conserved role in the control of neural A-P patterning via *Hox*-dependent mechanisms (Bel-Vialar et al., 2002; Charite et al., 1998; Isaacs et al., 1998; Shimizu et al., 2006). Cdx proteins are now also known to regulate neural A-P

patterning via *Hox*-independent mechanisms in different species (Skromne et al., 2007; Sturgeon et al., 2011). More recently, *Cdx1-Cdx2* double knockout mice were generated and revealed a novel redundant role for Cdx members in the control of NT closure (Savory et al., 2011a). This work showed that Cdx proteins regulate the planar cell polarity gene *Ptk7* and further suggested that Cdx members are involved in the regulation of convergent extension movements in the caudal embryo. This analysis involved a conditional mutagenesis approach to circumvent the peri-implantation lethality associated with the *Cdx2* null mutation, via a *CMV-ß-actin-Cre-ERT2* transgene and a floxed allele of *Cdx2.* Thus, the Cdx function was lost in all three germ layers and it is uncertain whether this novel Cdx role in neurulation is tissue-autonomous, -non-autonomous or due to a combination of both. *Ptk7* loss-of-function in *Xenopus* appears to affect convergent extension of the neuroectoderm (Wehner et al., 2011) but this has not been reported in the mouse. Indeed, although $Ptk7^{-/-}$ mouse mutants have been shown to have defective convergent extension movements in the mesoderm, an analysis of the neuroectoderm has not been reported (Yen et al., 2009). Therefore, more detailed analysis of *Cdx1-Cdx2* double knockout animals and conditional approaches involving tissue-specific loss-of-function will be required to clarify the Cdx-dependent processes involved in NT formation.

On the other hand, our work now suggests that Cdx proteins may impact on caudal neuroectoderm development in a tissue-autonomous manner, at least via the regulation of *Pax3* expression. As evidenced by the severe NC and NT defects observed in $Pax3^{Sp/Sp}$ mutants, *Pax3* plays a crucial role in the neuroectoderm (Auerbach, 1954; Li et al., 1999). Pax3 is important for NCC induction and analysis of *Pax3*-deficient embryos has indicated that NC defects are due to a marked reduction in the number of NCC that emigrate from the NT at cranial levels and a progressive complete loss of NCC at more caudal levels (Conway et al., 2000; Epstein et al., 2000; Olaopa et al., 2011). This progressive increase in the severity of NC defects along the A-P axis reflects the cranial co-expression of the functionally

redundant *Pax7* (Mansouri et al., 1996). Pax3 is also important for NT closure, being required for the survival of dorsal progenitors via downregulation of p53 activity (Pani et al., 2002; Wang et al., 2011). Thus, by acting upstream of *Pax3,* the canonical Wnt-Cdx pathway might control cell specification and maintenance of progenitor populations required for proper NC and NT development.

A role for Cdx proteins in NC development has not been formally reported in any species. This is most likely because such a role is masked by functional overlap and/or the presence of very severe posterior truncation phenotypes in *Cdx* compound mutants. However, several observations are in agreement with a role for Cdx proteins in NC development as well as the conservation of this role through evolution. In the mouse, an analysis of NCC in *Cdx1-Cdx2* double knockouts has not been reported but $Cdx1^{-/-}Cdx2^{+/-}$ mutants are known to display abnormal and fused dorsal root ganglia (van den Akker et al., 2002). In zebrafish, *Cdx* loss-of-function leads to a reduced number of Rohon-Beard cells (which share a common precursor with NCC) and absence of NC-derived spinal nerve roots (Epperlein et al., 2007; Skromne et al., 2007). In ascidian, *Cdx* loss-of-function results in absence of pigment cells which are derived from NC-like cells (Jeffery et al., 2004; Mita and Fujiwara, 2007). Future work focusing on a more detailed characterization of compound mutants or tissue-specific loss-of-function studies should help validate these observations and confirm a role for Cdx proteins in NC development.

1.6 Conclusion

In conclusion, our work suggests that Cdx proteins occupy a strategic position between canonical Wnt signals and *Pax3* at the beginning of the gene regulatory cascade controlling NCC development. Since *Cdx* genes are not expressed in the anterior neuroectoderm, the Wnt-Cdx pathway cannot impact on cranial NC induction. Therefore, our data are in accordance with the general idea that NCC are intrinsically different along the A-P axis (Abzhanov et al., 2003; Le Douarin et al.,

2004; Lwigale et al., 2004; Thibaudeau et al., 1998) and strongly suggest that these differences are already in place during induction of NCC.

1.7 Acknowledgements

The authors thank Denis Flipo (Université du Québec à Montréal) for the FACS analyses as well as Qinzhang Zhu and Li Lian (Institute de recherches cliniques de Montréal) for microinjections. Jonathan Epstein (Penn Cardiovascular Institute) is thanked for the *Pax3* probe, James B. Jaynes (Thomas Jefferson University) is thanked for the *Engrailed* cDNA and David Lohnes (University of Ottawa) is thanked for the Cdx antibodies as well as *Cdx1*-null mice. This work was supported by grants from the Canadian Institute for Health Research (CIHR grant number DCO190GP and IHD-94366) as well as from the Banting Research Foundation. OSF holds an Alexander-Graham-Bell scholarship from the Natural Science and Engineering Research Council (NSERC) of Canada, BC and TDS were supported by UQAM scholarships and IT was supported by a NSERC undergraduate research award. NP is a Fonds de la Recherche en Santé du Québec (FRSQ) Jr1 scholar.

1.8 Figures

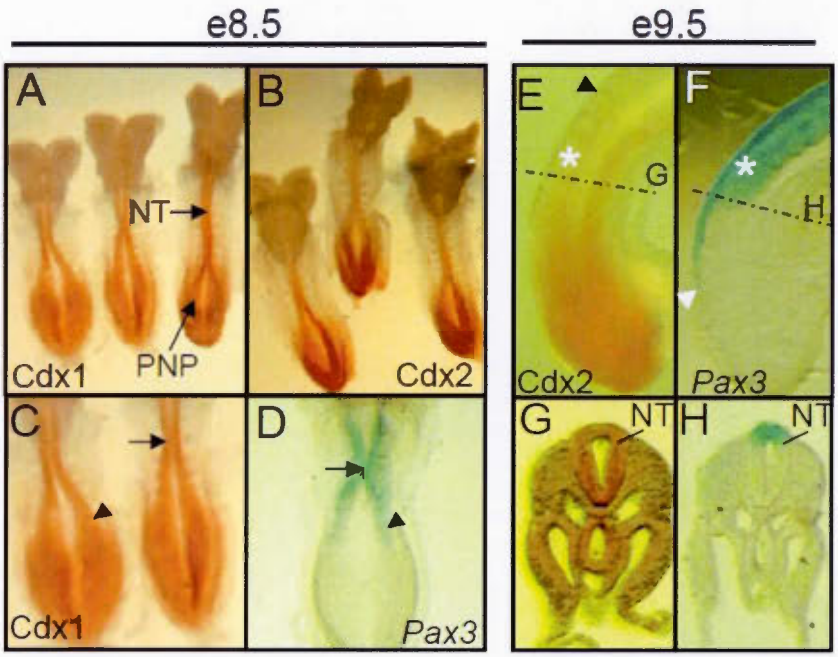


Figure 1.1 *Cdx* members and *Pax3* are co-expressed in the caudal neuroectoderm during the early steps of NT and NCC formation.

(A,B) Whole-mount immunohistochemistry showing Cdx1 *(A)* and Cdx2 *(B)* protein distribution in e8.5 embryos. Note the overlap in the posterior neural plate (PNP) and neural tube (NT). *(C-D)* Higher magnification view of Cdx1 protein distribution and comparison with *Pax3* gene expression domain as detected by whole-mount *in situ* hybridization. Note the overlap at the site of initiation of NT closure (arrow) and at the lateral PNP (arrowhead). *(E-H)* Comparison of Cdx2 (E, G) protein distribution with *Pax3* (F, H) gene expression domain in the caudal end of e9.5 embryos (lateral views). For comparative purposes, the location of the last formed somite is indicated by an asterisk. The black arrowhead in E indicates the anterior limit of Cdx2 expression; note the overlap with the posterior domain of *Pax3* gene expression (white arrowhead in F). The dotted lines in E-F indicate the level at which the transverse sections shown in G-H were cut. The Cdx2 protein is widely detected in the tailbud region (E) but this pattern becomes more restricted anteriorly, with notable strong detection in the whole NT (G). Note the overlapping detection of Cdx2 protein (G) and *Pax3* mRNA (H) in the posterior NT.

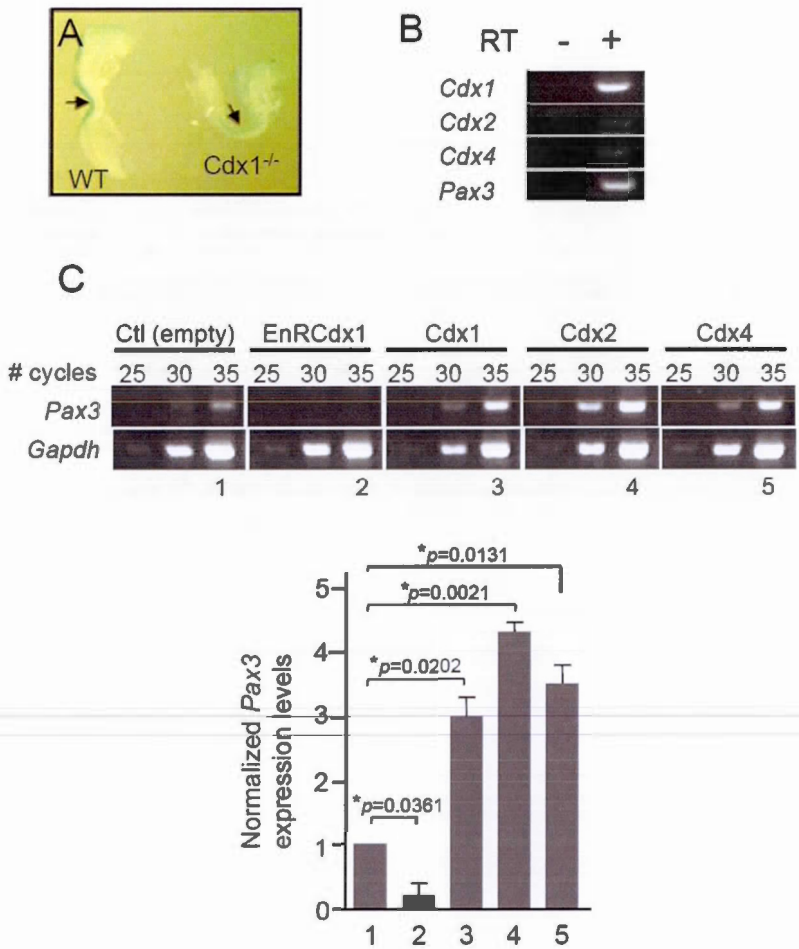


Figure 1.2 Regulation of *Pax3* expression by Cdx proteins.

(A) Whole-mount *in situ* hybridization analysis of *Pax3* expression in 6-somite (e8.5) wild-type (left) and *Cdx1*-null (right) mouse embryos. Embryos were processed and stained in parallel. Slightly less *Pax3* transcripts are detected in the dorsal NT (arrow) of the *Cdx1*-null embryo. *(B)* RT-PCR analysis showing co-expression of *Pax3* with the *Cdx* genes in N2a cells. ***(C)*** Semi-quantitative RT-PCR analysis showing alteration of endogenous *Pax3* expression levels upon modulation of Cdx activity in N2a cells. Prior to RT-PCR analysis, cells were transiently transfected with the indicated expression vector (GFP is co-expressed as a bicistronic transcripts, allowing FACS-mediated recovery of transfected cells). *Pax3* expression levels are normalized against *Gapdh* expression. Note that *Pax3* expression is drastically reduced in cells expressing EnRCdx1 whereas it is significantly increased in cells overexpressing a Cdx member. Numbers 1-5 represent the samples for which signal intensity was assessed by densitometry and results used

to generate the graph on the bottom. Note that similar results were obtained from three independent experiments.

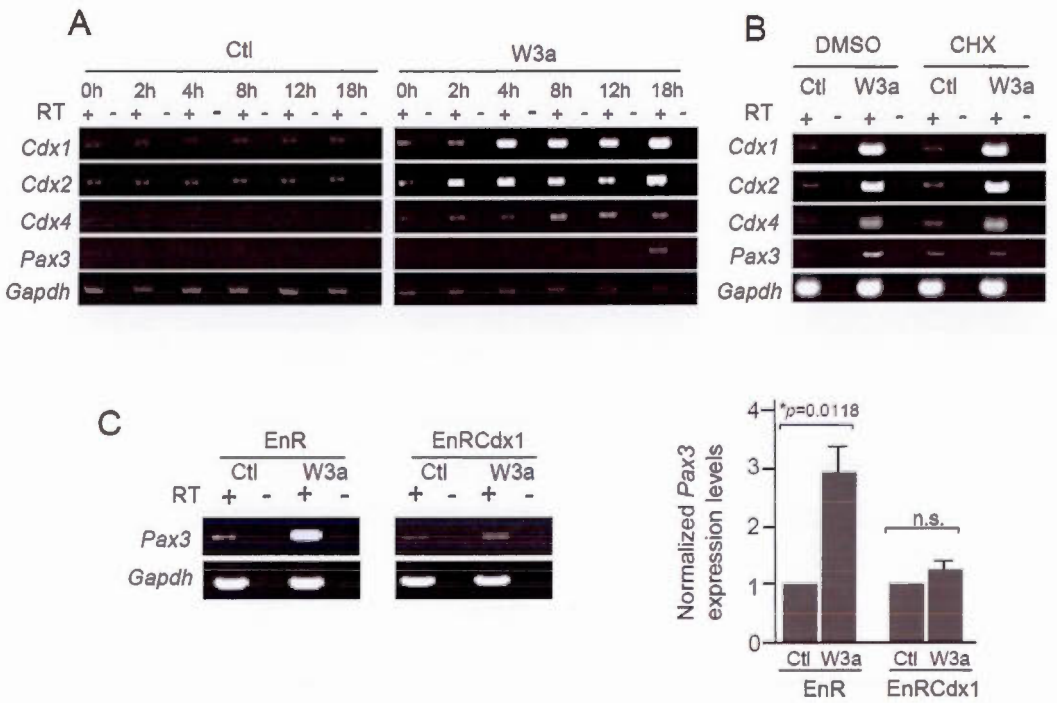


Figure 1.3 Regulation of *Pax3* expression by the Wnt-Cdx pathway.

(A-C) Analyses of P19 cells cultured in Wnt3a-conditioned medium (W3a) or control-conditioned medium (Ctl). After treatment, expression of *Cdx1, Cdx2, Cdx4* and *Pax3* was assessed by RT-PCR. *Gapdh* was used as a loading control. Note that expression of *Cdx1, Cdx2, Cdx4* and *Pax3* is specifically induced by treatment with W3a whereas no induction is seen with Ctl. Shown are representative results obtained from three independent experiments *(A)* Time course analysis in cells incubated for the indicated time with Ctl (left panels) or W3a (right panel); note that expression of all *Cdx* members is induced by W3a much earlier than *Pax3*. *(B)* Dependence of *de novo* protein synthesis for *Pax3* induction. Cells were pre-treated for 30 min with a vehicle (DMSO) or 30 µg/mL of cycloheximide (CHX) and then cultured for 24h in Ctl or W3a in presence of 1µg/ml of CHX. Note that *Pax3* induction is affected by CHX treatment while expression of *Cdx1*, *Cdx2* and *Cdx4* is not. *(C)* Semi-quantitative RT-PCR analysis showing alteration of endogenous *Pax3* expression level in the presence of the Cdx dominant negative fusion protein EnRCdx1. Prior to RT-PCR analysis, cells were transiently transfected with EnR (control) or EnRCdx1 expression vector which co-expresses GFP as a bicistronic transcript. Approximately 30 hours after transfection cells were cultured for another 24 hours in the presence or absence of W3a. GFP-positive cells were then recovered by FACS, total RNA extracted and endogenous *Pax3* expression assessed by RT-PCR. Note that *Pax3* induction is reduced in cells

expressing EnRCdx1 but not in cells expressing EnR. *Pax3* expression levels were normalized against *Gapdh* expression.

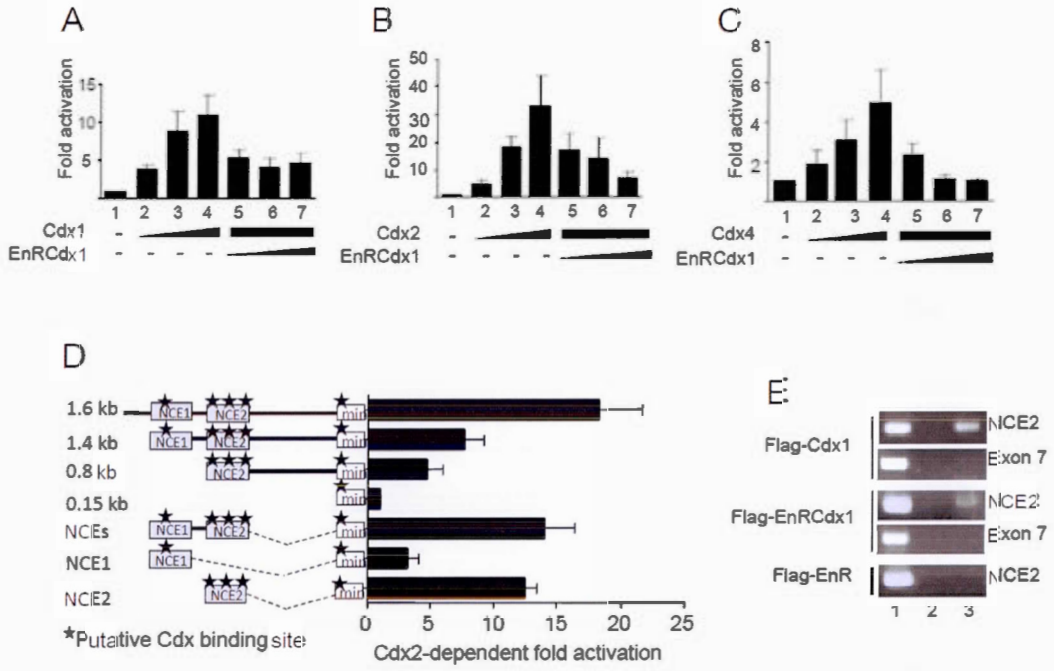


Figure 1.4 Identification of Cdx responsive regions in the *Pax3* proximal promoter.

(A-D) Co-transfection assays in N2a cells using Luciferase reporter constructs driven by 5'-flanking sequences of *Pax3*. The results (mean±SEM of 4-5 independent experiments performed in triplicate) are expressed as fold induction compared to the relevant reporter vector alone. *(A-C)* Cdx-dependent transactivation of a *Pax3p1.6kb*-luciferase construct. Note that addition of (A) Cdx1, (B) Cdx2 or (C) Cdx4 expression construct results in strong dose-dependent transactivation (lanes 2-4), which is repressed by increasing amounts of EnRCdx1 (lanes 5-7). *(D)* Detection of Cdx responsive regions in the *Pax3* promoter via co-transfection assays in N2a cells. Luciferase reporter constructs consisting of the *Pax3* 150bp minimal promoter (min) with or without various lengths or regions of the *Pax3* promoter were assayed for Cdx2 transactivation. Potential Cdx binding sites are indicated by black stars. Note that Cdx2 responsiveness correlates with the presence of potential Cdx binding sites, which are located in previously identified neural crest enhancers (NCE), and especially with the three putative binding sites located in NCE2. *(E)* Chromatin immunoprecipitation assays in N2a cells showing the presence of Cdx1 and EnRCdx1 proteins on the endogenous *Pax3* promoter. Primers flanking *Pax3* NCE2 or exon 7 (as a negative control) were used to amplify anti-Flag immunoprecipitated DNA from Flag-Cdx1 (upper panels), Flag-EnRCdx1 (middle panels) or Flag-EnR (lower panel) transfected cells. 1, input; 2, pre-immune serum IP; 3, anti-Flag IP. Note in lane 3 that a PCR product for NCE2 is obtained from Flag-Cdx1 and Flag-EnRCdx1 transfected cells but not from the negative control Flag-EnR transfected cells.

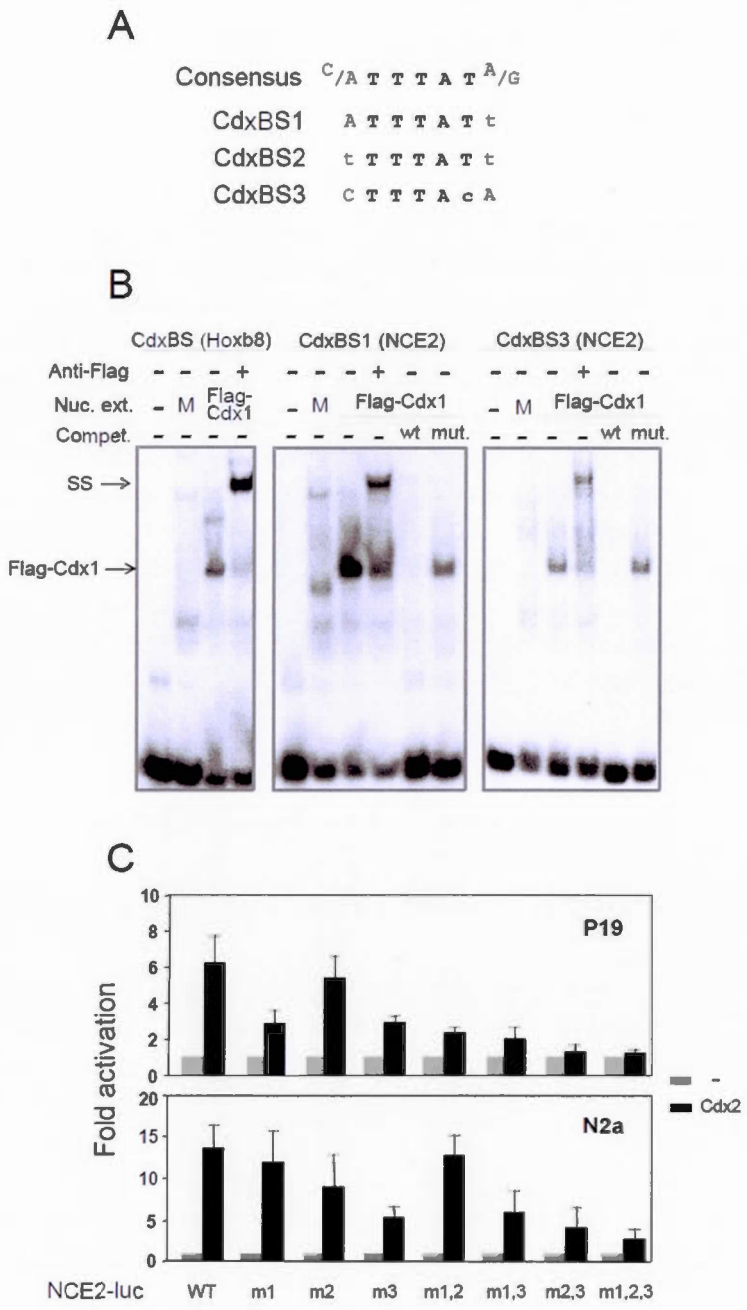
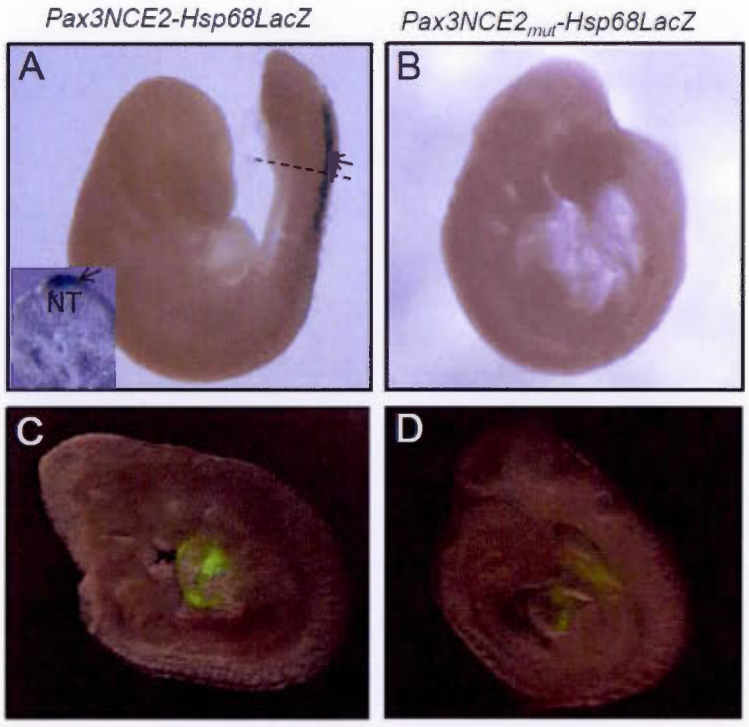


Figure 1.5 Identification and characterization of Cdx binding sites in *Pax3* NCE2.

(A) Sequence comparisons of all three putative Cdx-binding sites (CdxBS1, CdxBS2 and CdxBS3) relative to a consensus CdxBS sequence. Mismatches are denoted by lowercase

lettering. *(B)* Analysis of Cdx1 binding to CdxBS1 and CdxBS3 via electrophoretic mobility shift assay. All *in vitro* binding reactions were performed in parallel under identical conditions. M, mock. The presence of Cdx1 in the shifted bands was confirmed by addition of 1 μg of anti-Flag antibody to the binding reaction mix. Cdx1 binding and anti-Flag supershifts (SS) are indicated by arrows. A probe containing a consensus CdxBS present in the *Hoxb8* promoter was used as a positive control. Specificity of Cdx1 binding was assessed by pre-incubation of Flag-Cdx1 containing nuclear extracts with a 100 fold excess of wild-type (wt) or mutated (mut) cold probes. Note that pre-incubation with wt unlabeled probes leads to inhibition of Cdx1 binding to the radiolabeled probes, whereas pre-incubation with mut probes did not. *(C)* Impact of CdxBS mutation on Cdx2-mediated activation of the NCE2 reporter in cell culture. Wild type (wt) or CdxBS mutant versions of a *Pax3*NCE2-luciferase reporter were generated and assessed for Cdx2-mediated transactivation in P19 and N2a cells. Cells were transiently transfected with the NCE2-Luc reporter alone or with a *Cdx2* expression vector. The results (mean±SEM of 7 to 8 independent experiments performed in triplicate) are expressed as fold induction compared to the relevant reporter vector alone. m1, m2 or m3 denote *Pax3*NCE2-luciferase reporter constructs containing point mutations in the CdxBS1, CdxBS2 or CdxBS3 respectively; m1,2, m1,3 or m2,3 indicate double mutations of CdxBS whereas m1,2,3 denote the triple mutation. Note that concomitant mutation of all three CdxBS is required to almost completely abolish the transactivation of *Pax3* NCE2 by Cdx2.



E

	Number of transgenics	Dorsal NT expression
NCE2	6	4
$NCE2_{mut}$	4	0

Figure 1.6 Cdx binding sites are required for the activity of a *Pax3*NCE2-lacZ reporter in transgenic embryos.

(A-B) Whole mount ß-Galactosidase staining performed in e9.5 transgenic embryos generated from (A) the wild type (*Pax3*NCE2-*lacZ*) or (B) mutated (*Pax3*$NCE2_{mut}$-*lacZ*) transgene, the later carrying specific point mutations in all three CdxBS of NCE2. The dotted line in (A) indicates where the transverse section shown in the inset was cut. *(C-D)* Direct GFP fluorescence detection of a *Gata4p-GFP* transgene co-injected with (C) the wild type or (D) mutated transgene. Note that staining is detected in the dorsal NT (arrow) of transgenic embryos generated from the wild-type construct (A). Mutation of all CdxBS abrogates ß-Galactosidase activity in the dorsal NT (B) while detection of *Gata4p-GFP* is not affected (D). *(E)* Table indicating the total number of transgenic embryos obtained for each construct as well as the number of embryos exhibiting ß-Galactosidase activity in the dorsal NT.

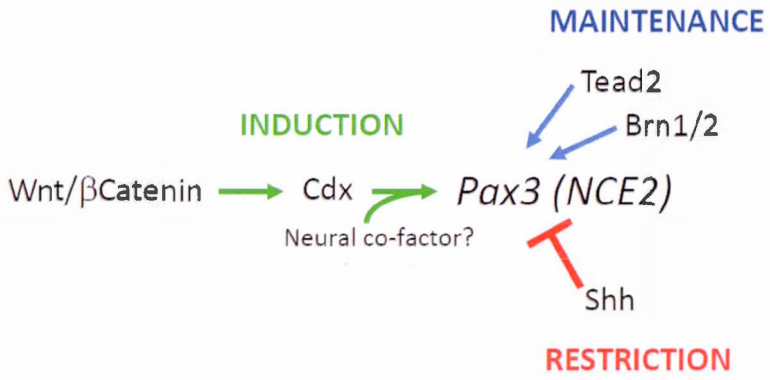


Figure 1.7 Control of *Pax3* expression in the caudal neuroectoderm via NCE2.

Induction of *Pax3* expression in the posterior neural plate is controlled by the Wnt-Cdx pathway. Expression in the closed neural tube is later maintained by the activity of Tead2 as well as Brn1/2 transcription factors. Restriction of *Pax3* expression at the lateral neural plate and dorsal neural tube is ensured by repressive Shh signals emerging from the node, notochord and floorplate. An unknown neural-specific Cdx co-factor might also be involved in the spatial restriction of NCE2 activity.

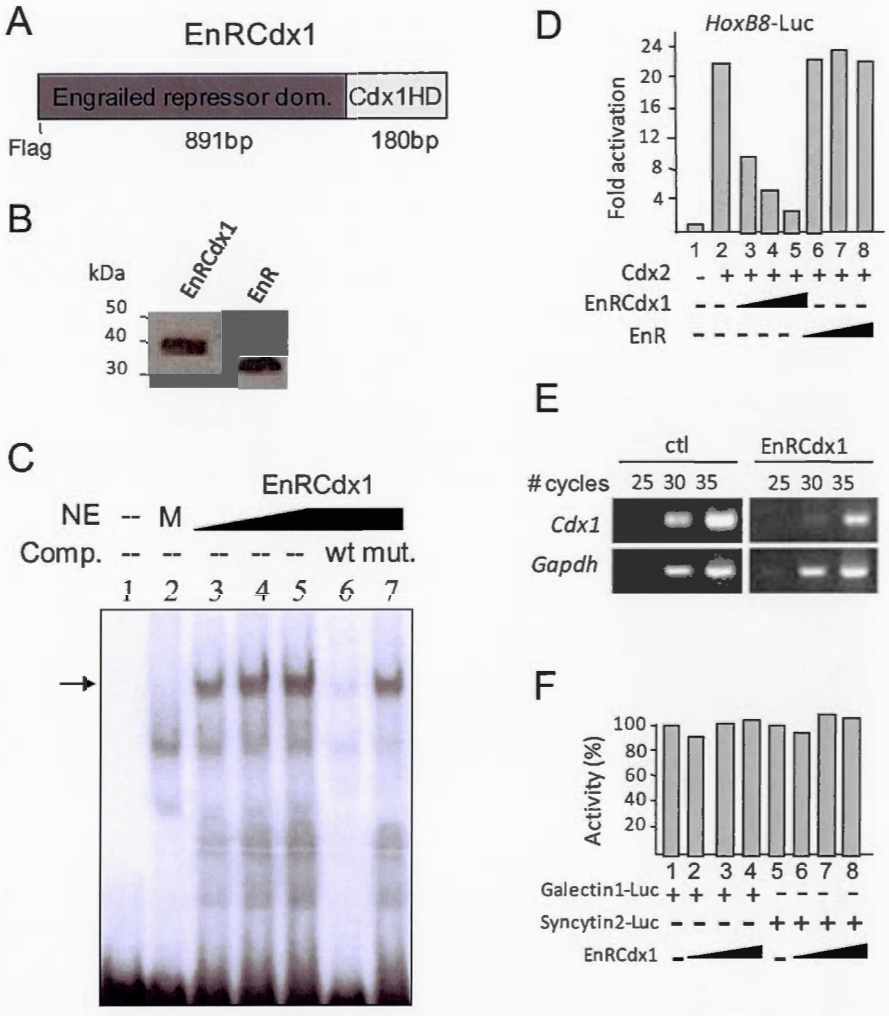


Figure 1.8 Characterization of the Cdx dominant negative fusion protein.

(A) The Cdx dominant negative fusion protein consists of the repressor domain of Engrailed (EnR) and the homeodomain of Cdx1. *(B)* Anti-Flag western blot showing detection of EnRCdx1 and EnR proteins in Cos7 cells transfected with respective expression vectors. *(C)* Electrophoretic mobility shift assay using a consensus Cdx binding site as probe and nuclear extracts (NE) from Cos7 cells transfected or not with an EnRCdx1 expressing vector: 1-free probe; 2-mock NE; 3-5-increasing amount of EnRCdx1 NE; 6-7-maximum amount of EnRCdx1 NE and competition with wt (6) or mutated (7) cold probes. EnRCdx1 binding is indicated by an arrow. *(D)* Co-transfection assay in P19 cells using a *Hoxb8* reporter construct bearing multiple Cdx binding sites. Cdx2 expression results in strong transactivation (lane2). Co-transfection of Cdx2 and increasing amounts of EnRCdx1 expression construct results in dosedependent repression (lanes 3-5) whereas expression of

the EnR domain without the Cdx1 homeodomain does not affect Cdx2-mediated transactivation (lanes 6-8). *(E)* Semi-quantitative RT-PCR analysis showing repression of endogenous Cdx1 expression in N2a cells transfected with EnRCdx1 expression vector. Note that *Cdx1* is known to autoregulate via binding of Cdx1-Lef1 complexes to Lef/Tcf binding sites in the *Cdx1* promoter (46). *Cdx1* expression levels are normalized against *Gapdh* expression. Number above each lane represents the number of PCR cycles. *(F)* Co-transfection assays in N2a cells using luciferase reporter plasmids driven by Galectine1 or Syncytin2 promoter (provided by B. Barbeau, UQAM) to confirm the specificity of EnRCdx1. Increasing amounts of EnRCdx1 do not affect the expression of these non-Cdx target reporters.

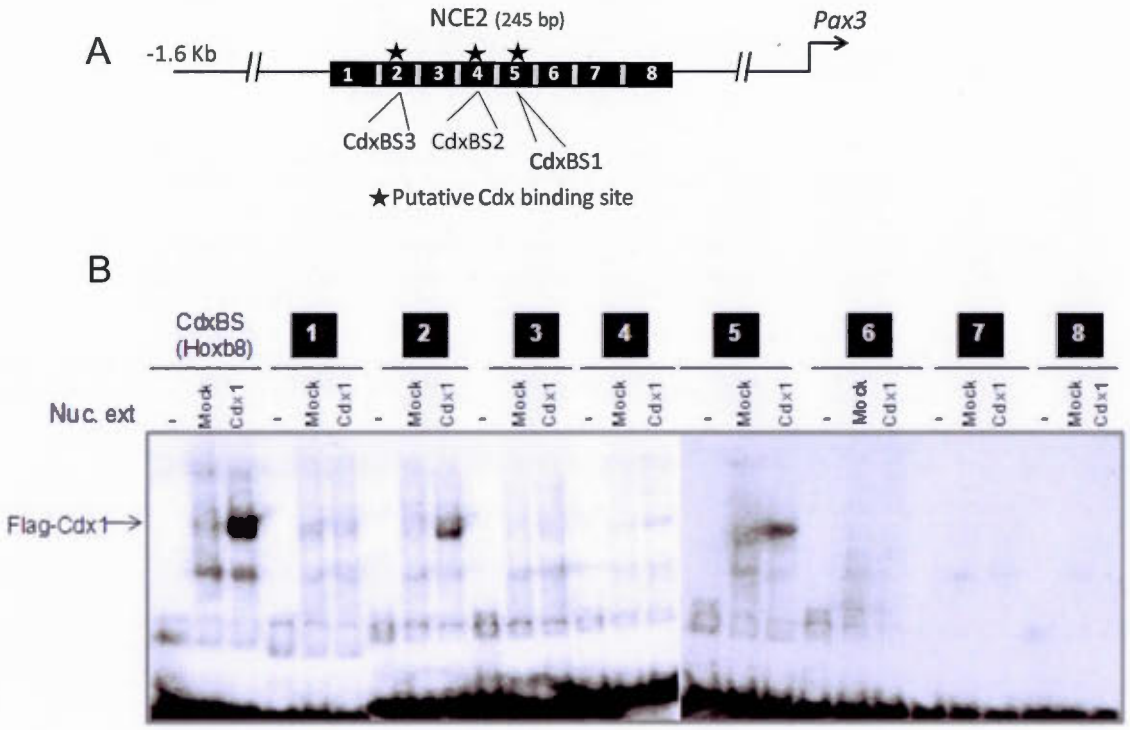


Figure 1.9 Identification of Cdx binding sites in the *Pax3* NCE2.

(A) Schematic representation of the *Pax3* NCE2 showing the relative location of all three putative Cdx binding sites (CdxBS1, CdxBS2 and CdxBS3) as well as the 8 overlapping oligonucleotides (denoted 1-8) used to scan the NCE2 for Cdx binding sites in gel shift assay. ***(B)*** Identification of Cdx binding sites in the *Pax3* NCE2 by electrophoretic mobility shift assay (EMSA). A probe containing a consensus CdxBS present in the *Hoxb8* promoter was used as a positive control. Each probe was incubated or not (-) with nuclear extracts from mock or Flag-Cdx1 (Cdx1) transfected Cos7 cells. Note that strong Cdx binding is only observed with oligonucleotide probes 2 and 5 bearing respectively the CdxBS3 and CdxBS1.

Table 1.1 Oligonucleotides used to identify CdxBS in the *Pax3* NCE2 by EMSA.

Probe	Sequence
1	5'-AGGCTAGGCACAATGGTACCTTCTCT*AAGGACAGAC*-3'
2	5'-*AAGGACAGAC*AGTCTTTACAACACTCCTGG*CGTCATATCC*-3'
3	5'-*CGTCATATCC*TGCTGGGGACACTTCAGCTC*CTAGCCAAGA*-3'
4	5'-*CTAGCCAAGA*CGTTGCTTCTTTTATTTTTC*CAGCAGTTTA*-3'
5	5'-*CAGCAGTTTA*GTCTGAATGCCATAATAAATTCCTGAGAAC-3'
6	5'-*TCCTGAGAAC*AAACGCTGCACCCGGGCAAA*ACCTCAACAT*-3'
7	5'-*ACCTCAACAT*ATAGATGCAAGTGCATCGGG*GATGAATGTG*-3'
8	5'-*GATGAATGTG*TACGTGGAGATTTAAAGTCCCCGCTTCTCA-3'
Consensus CdxBS Hoxb8	5'-GACCGCTATAAAAGTTTATAGGGTATAAATTTCTGA-3'
CdxBS1mut	5'-CAGCAGTTTAGTCTGAATGCCATAATACCTTCCTGAGAAC-3'
CdxBS3mut	5'-AAGGACAGACAGTCTCGACAACACTCCTGGCGTCATATCC-3'

Putative, consensus or mutated CdxBS sequences are underlined, overlapping sequences are in italics and point mutations are denoted in bold

CHAPTER II

Induction and dorsal restriction of *Paired-box 3* (*Pax3*) gene expression in the caudal neuroectoderm is mediated by integration of multiple pathways on a short neural crest enhancer

(Published in BBA-Gene Regulatory Mechanisms, 2014, vol 1839, pages 546-58)

Oraly Sanchez-Ferras, Guillaume Bernas, Emilie Laberge-Perrault and Nicolas Pilon.

Molecular Genetics of Development Laboratory, Department of Biological Sciences and BioMed Research Center, Faculty of Sciences, University of Quebec at Montreal (UQAM)

To whom correspondence should be addressed: Nicolas Pilon, Email: pilon.nicolas@uqam.ca

Author contribution:

Oraly Sanchez-Ferras: conception, design, acquisition, analysis and interpretation of data, drafting and revision of the article; (experiments: all figures in the paper). *Guillaume Bernas*: acquisition of data: (experiments: production of Glutathione *S*-transferase fusion proteins; immunofluorescence analysis of Cdx2 and Pax3 expression in NT cryosections, co-transfection analyses in P19 and N2a cells). *Emilie Laberge-Perrault:* preliminary experiments. *Nicolas Pilon:* supervision of the work and revision of the manuscript.

2.1 Summary

Pax3 encodes a paired-box transcription factor with key roles in neural crest and neural tube ontogenesis. Robust control of *Pax3* neural expression is ensured by two redundant sets of cis-regulatory modules (CRMs) that integrate anterior-posterior (such as Wnt-ßCatenin signaling) as well as dorsal-ventral (such as Shh-Gli signaling) instructive cues. In previous work, we sought to characterize the Wnt-mediated regulation of *Pax3* expression and identified the Cdx transcription factors (Cdx1/2/4) as critical intermediates in this process. We identified the neural crest enhancer-2 (NCE2) from the 5'-flanking region of *Pax3* as a Cdx-dependent CRM that recapitulates the restricted expression of *Pax3* in the mouse caudal neuroectoderm. While this is consistent with a key role in relaying the inductive signal from posteriorizing Wnt ligands, the broad expression of Cdx proteins in the tailbud region is not consistent with the restricted activity of NCE2. This implies that other positive and/or negative inputs are required and, here, we report a novel role for the transcription factor Zic2 in this regulation. Our data strongly suggests that Zic2 is involved in the induction (as a direct *Pax3*NCE2 activator and Cdx neural cofactor) as well as the maintenance of *Pax3* dorsal restriction (as a target of the ventral Shh repressive input). We also provide evidence that the inductive Cdx-Zic2 interaction is integrated on NCE2 with a positive input from the neural-specific transcription factor Sox2. Altogether, our data provide important mechanistic insights into the coordinated integration of different signaling pathways on a short *Pax3* CRM.

2.2 Introduction

Building an embryo from an initial population of equivalent cells requires precise spatiotemporal control of gene expression. Information to do this comes from just a few conserved signaling pathways, is transmitted by DNA binding proteins and interpreted at the cis-regulatory level on evolutionarily conserved genomic sequences. Redundant and different operating enhancers may exist to refine and protect

expression of developmental genes from fluctuations in these signals or mutations in the genome. Subsequently, the output of gene expression and gene regulatory interactions provide the memory to maintain established expression patterns in the absence of signaling inputs. In this regard, a lot of work has been done to understand the molecular mechanisms of neural gene expression during establishment of the anterior-posterior (A-P) as well as the dorsal-ventral (D-V) axes. Nevertheless, how both A-P and D-V signaling inputs are coordinately integrated at the cis-regulatory level is still poorly understood.

Pax3/7 (Paired box 3 and 7) and *Zic* (Zinc finger protein of the cerebellum) family members (*Zic1-5*) encode transcription factors that exhibit overlapping expression domains in the neuroectoderm along both the A-P and the D-V axis. During neurulation, expression of these genes is similarly restricted to pre-migratory neural crest cells (NCC) and dorsal neural tube (NT) (Aruga, 2004; Auerbach, 1954; Epstein et al., 1991; Gaston-Massuet et al., 2005; Goulding et al., 1991; Inoue et al., 2004; Li et al., 1999; Mansouri et al., 1996; Nagai et al., 1997). However, such overlap is less extensive along the AP axis and most especially in the caudal embryo where only *Pax3* and the *Zic2/Zic5* gene pair are expressed in the posterior neural plate (PNP) (Goulding et al., 1991; Merzdorf, 2007). Consistent with their wide expression pattern and key developmental role, loss-of-function mutations of *Pax3* and *Zic2* – as seen for example in the Splotch (*Pax3*) and Kumba (*Zic2*) mouse mutants – causes severe and similar NT and NCC defects affecting the entire A-P axis such as spina bifida, craniofacial malformations, absence of dorsal root ganglia and pigmentary anomalies (Auerbach, 1954; Elms et al., 2003; Li et al., 1999; Merzdorf, 2007; Nagai et al., 2000).

Work performed in several vertebrate species has revealed that A-P instructive cues from Wnt (Wingless and Int-1 related) and FGF (fibroblast growth factor) pathways as well as D-V instructive cues from BMP (bone morphogenetic protein) and Shh

(Sonic Hedgehog) pathways are all involved in the induction and dorsal restriction of *Pax3/7* and *Zic* members (Aruga, 2004; Bang et al., 1999; Bang et al., 1997; de Croze et al., 2011; Ericson et al., 1996; Garnett et al., 2012; Goodrich et al., 1997; Goulding et al., 1993; Liem et al., 1995; Monsoro-Burq et al., 2005; Taneyhill and Bronner-Fraser, 2005). These studies notably point to a critical role for posteriorizing canonical Wnt signaling and intermediate levels of BMP molecules during induction in the neural plate (Bang et al., 1999; de Croze et al., 2011; Garnett et al., 2012; Monsoro-Burq et al., 2003; Sato et al., 2005; Taneyhill and Bronner-Fraser, 2005) whereas opposing gradients of dorsal BMP and ventral Shh signaling are subsequently implicated in the maintenance and dorsal restriction in the closed NT (Aruga et al., 2002b; Goodrich et al., 1997; Goulding et al., 1993; Liem et al., 1995). Although the general role of these pathways is well accepted, some species-specific variations are also expected regarding their relative importance. In the case of *Pax3*, this is well exemplified by the comparison of its posterior expression domain between chick and mouse embryos. Indeed, in chick embryos, *Pax3* expression is initially induced in the whole PNP before becoming restricted to the dorsal NT whereas in mouse embryos, *Pax3* expression is already restricted to the lateral borders of the PNP during the induction phase (Bang et al., 1997; Sanchez-Ferras et al., 2012).

Multiple evolutionary conserved cis-regulatory modules (CRMs) have been identified for *Pax3*. These CRMs are clustered in two areas of the *Pax3* locus: one in the 5'-flanking region and the other in intron-4 (Degenhardt et al., 2010; Milewski et al., 2004; Natoli et al., 1997). The 5'-flanking region, named Neural Crest Enhancer (NCE), is located within the 1.6 kb proximal promoter and is subdivided in two short CRMs of approximately 250 bp named NCE1 and NCE2 (Milewski et al., 2004; Natoli et al., 1997). The entire NCE is not only able to direct *Pax3* reporter expression in mouse NCC and dorsal NT along the hindbrain and trunk region, but also drive enough functional expression levels of *Pax3* to rescue the NT and NCC defects observed in *Pax3* Splotch mice (Li et al., 1999). Interestingly, targeted

deletion studies in the mouse have suggested that NCE acts redundantly with a second evolutionary conserved region (ECR2) located in the 4th intron (Degenhardt et al., 2010). In fact, more recent work using the zebrafish as a model has demonstrated that the *Pax3* intron-4 contains at least two CRMs that appears to exhibit complementary activities in order to recapitulate the induction, dorsal restriction and maintenance of *Pax3* neural expression (Garnett et al., 2012; Moore et al., 2013). Given that the effect of their deletion has not been documented so far, the requirement of any of these intron-4 CRMs for *Pax3* expression as well as their relative importance over the 5'-flanking NCE is currently unknown.

We have previously demonstrated that the NCE2 region alone is able to recapitulate both the induction and dorsal restriction of *Pax3* expression in the caudal NCC and NT (Sanchez-Ferras et al., 2012), suggesting that this CRM is well suited for analyzing the coordinated integration of both AP and DV instructive cues. In this regard, we have already demonstrated that activity of this enhancer depends on a positive input from caudal-related homeobox (Cdx) transcription factors downstream of Wnt/ßCatenin signaling (Sanchez-Ferras et al., 2012). Here we further show that, in addition to Cdx, Zic2 also directly regulates murine *Pax3* expression and acts as a Cdx neural cofactor. Importantly, we show that the NCE2 region integrates positive inputs from caudal Cdx, dorsal Zic2 as well as neural Sox2 transcription factors. Furthermore, we provide evidence for a putative role of Zic2/5 as mediators of the Shh-induced repressive input involved in the dorsal restriction of *Pax3* expression. Taken together with previous descriptions of other functional binding sites (e.g. Brn1 and Tead2) within NCE2 (Milewski et al., 2004; Pruitt et al., 2004), our data strongly suggest that this short CRM behaves as a "super-enhancer" (Whyte et al., 2013) that mediates the spatiotemporal induction and dorsal restriction of *Pax3* expression in the mouse caudal neuroectoderm.

2.3 Materials and Methods

2.3.1 Ethics Statement

Experiments involving mice were performed following Canadian Council of Animal Care (CCAC) guidelines for the care and manipulation of animals used in medical research. Protocols involving the manipulation of animals were approved by the institutional ethics committee of the University of Quebec at Montreal (comité institutionnel de protection des animaux (CIPA)); Reference number 0513-C1-648-0514).

2.3.2 Plasmid constructs and site-directed mutagenesis

The *Pax3* cDNA vector pBH3.2 (Goulding et al., 1991) was kindly provided by J. Epstein. Expression vectors for FLAG-tagged Cdx1 and GST-Cdx1 fusion proteins have been described previously (Beland et al., 2004). Expression vectors for FLAG-tagged full-length and deletion mutant ZIC2 proteins were a generous gift from S. Tejpar (Pourebrahim et al., 2011). HA-tagged Zic2 expression vector (pcDNA3-$_{HA}$Zic2) was kindly provided by J. Aruga (Koyabu et al., 2001). Myc-tagged Sox2 expression vector (pcDNA3.1-$_{Myc}$Sox2) was a gift from M. Bani-Yaghoub (Bani-Yaghoub et al., 2006). $_{FLAG}$Cdx1 and $_{HA}$Zic2 expression vectors co-expressing GFP were generated by subcloning the respective cDNA into the pIRES2-EGFP vector (Clontech). Wild type (wt) or CdxBS mutant versions of a *Pax3*NCE2-luciferase reporter constructs were as previously described (Sanchez-Ferras et al., 2012). For generating Zic binding site (BS) mutant versions, point mutations were introduced into the ZicBS of NCE2 by using the Quick Change Multisite-directed mutagenesis kit (Stratagene) in accordance with manufacturer's instructions. Sequence of the oligonucleotide used for site-directed mutagenesis of the ZicBS was: 5'-CGTCATATCCTGCTaaGGACACTTCAGCTCCTAGCCAAGA-3' (with ZicBS underlined and point mutations indicated by lower case letters).

2.3.3 In situ hybridization and immunofluorescence analyses

FVB mouse embryos were obtained from timed pregnancies, noon of the day on which a vaginal plug was detected being considered as embryonic day (e) 0.5. Embryos to be compared were stage-matched in accordance with established criteria (Kaufman, 1992) and processed in parallel.

Whole-mount *in situ* hybridization of *Pax3* and *Zic2* mRNAs was performed using standard approach (Wilkinson, 1992). The *Pax3* probe was generated from the pBH3.2 plasmid while the *Zic2* probe was generated from the pcDNA3-$_{HA}$Zic2 plasmid. Transverse sections (100µm) were prepared using a vibrating blade microtome Microm HM 650V (Thermo Scientific) as previously described (Coutaud and Pilon, 2013b). Images were acquired with a Leica DFC 495 camera mounted on a Leica M205 FA microscope (Leica Microsystems).

Whole mount immunostaining of Nkx6.1 in e9.5 mouse embryos was performed using the protocol described by (Coutaud and Pilon, 2013b) and consecutive vibratome transverse sections of the tail were prepared as described above. Immunofluorescence analyses on NT cryosections were performed as previously described (Jeong and McMahon, 2005). Briefly, 20 µm frozen sections were prepared using a Leica CM1950 cryostat (Leica microsystems) and adjacent sections were used to compare the distribution of Cdx2 and Pax3 proteins. Sections were blocked with 10% fetal bovine serum, 0.1% TritonX-100 in PBS for 1 hour and then incubated at 4°C overnight with primary antibodies diluted in blocking solution. Nuclei were stained using DAPI (Molecular Probes) and slides were then incubated 1 hour at room temperature with secondary antibodies diluted in blocking solution. Antibodies and dilutions used were: mouse anti-Cdx2 1: 100 (Biogenex), mouse anti-Pax3 1:400 (R&D Systems), mouse IgG1 anti-NKX6.1 1:20 (Developmental Studies Hybridoma Bank) and Alexa-647 donkey anti-mouse 1:500 (Jackson

Immunoresearch). Images were taken with a Nikon A1 laser-scanning confocal microscope.

2.3.4 Chromatin immunoprecipitation (ChIP) assays

ChIP assays in Neuro2a (N2a) cells were performed using the M-Fast Chromatin immunoprecipitation kit (ZmTech Scientific) in accordance with manufacturer's instructions. For immunoprecipitation of chromatin from $_{HA}$Zic2- or $_{FLAG}$Cdx1-transfected N2a cells (10^6 cells), 1μg of rabbit anti-HA (Abcam) and mouse anti-FLAG M2 (Sigma) antibodies were respectively used. A normal rabbit IgG serum was used as a negative control of immunoprecipitation. PCR amplifications were performed using the Platinum Taq DNA polymerase (Life Technologies) and consisted of 35 cycles of 30s at 96°C, 30s at 60°C and 30s at 72°C. The primers used for this study were: *Pax3* NCE2 Forward (5'-GGCACAATGGTACCTTCTCTAAGG-3') and Reverse (5'-CCCTTCTGAGAAGCGGGGACTTTAAA-3'). PCR products were resolved on a 2 % agarose gel and sequence-confirmed.

2.3.5 Electrophoretic mobility shift assays (EMSA)

Zic2 and Cdx1 binding to *Pax3*NCE2 sequences was assessed essentially as previously described (Sanchez-Ferras et al., 2012). Briefly, binding reactions were carried out in a previously described Zic binding buffer (Mizugishi et al., 2001) with 8□g of nuclear extracts from mock-, $_{HA}$Zic2 or $_{FLAG}$Cdx1 transfected Cos7 cells. Supershift assay was performed using 1μg of rabbit anti-HA antibody (Abcam). Specificity of binding was assessed by competition with a 100-fold molar excess of unlabeled wild type or mutated cold probe. To test the interaction of Cdx1 and Zic2 in the presence of the target DNA, $_{HA}$Zic2 and $_{FLAG}$Cdx1 nuclear extracts were incubated for 30 min in Zic binding buffer at 4 °C before the addition of the radiolabelled probe (either *Pax3*NCE2 CdxBS1, CdxBS3 or ZicBS). The sense strand of each wild type double-stranded probe was previously described (Sanchez-Ferras et

al., 2012). The sense strand of a double-stranded probe harbouring the mutated *Pax3*NCE2 ZicBS was as described above for site-directed mutagenesis. The sense strand of a double stranded probe harbouring a previously described ZicBS (Mizugishi et al., 2001) (used as positive control) was: 5'-GATCCTGTGATTTTCGTCTTGGGTGGTCTCCCTCG-3' (with ZicBS underlined).

2.3.6 Western blot, co-immunoprecipitation and GST pull-down assays

Primary antibodies and dilutions used for western blotting were: mouse anti-FLAG 1:1000 (M2, Sigma), rabbit anti-HA 1:2000 (ab9110, Abcam), mouse anti-Myc 1:100 (in house 9E10 hybridoma) and rabbit anti-GAPDH 1:2500 (sc25778, Santa Cruz Biotech).

For co-immunoprecipitation, Cos7 cells were transfected with expression vectors for $_{FLAG}$Cdx1 and $_{HA}$Zic2 using Genejuice transfection reagent (Novagen) and harvested in immunoprecipitation buffer (25 mm Hepes, pH 7.2, 0.5% Nonidet P-40, 150 mm NaCl, supplemented with protease inhibitors). Immunoprecipitations and western blots were carried out as previously described (Beland et al., 2004).

For pull-down assays, Glutathione *S*-transferase (GST)-Cdx1, GST-Cdx1 homeodomain (GST-Cdx1Homeo) and GST-Cdx1 N-terminus (GST-Cdx1Nterm) fusion proteins were produced and purified with glutathione-agarose beads (Sigma) as previously described (Beland et al., 2004). FLAG-tagged full-length and deletion mutant ZIC2 proteins were *in vitro* synthesized using the T_NT T7 quick coupled transcription/translation system (Promega) and previously described plasmids (Pourebrahim et al., 2011). *In vitro* pull-down assays were performed as previously described (Beland et al., 2004) and interactions were revealed by western blotting using mouse anti-FLAG M2 antibody (Sigma). To test the interaction between ZIC2 and Cdx1 in the presence of NCE2 sequences, *in vitro* synthesized $_{FLAG}$ZIC2 proteins

were pre-incubated with increasing amounts of NCE2 DNA (50 or 200 ng) for 30 min in the same conditions as described for EMSA before adding GST or GST-Cdx1 homeodomain fusion proteins.

2.3.7 Cell culture, transfections and RT-PCR analyses

P19, Neuro2a (N2a) and Cos7 cell lines were propagated and transfected as previously described using Genejuice reagents (Novagen) in accordance with manufacturer's protocol (Sanchez-Ferras et al., 2012).

Luciferase reporter assays were performed essentially as previously described (Sanchez-Ferras et al., 2012). For Zic2 or Sox2 dose response assays, N2a ($8x10^4$ cells per well) and P19 ($3x10^4$ cells per well) were transfected with 100 ng of *Pax3*NCE2-Luciferase reporter construct alone or with increasing amounts of $_{HA}$Zic2 (10 ng to 40 ng) or $_{Myc}$Sox2 (10 ng to 75 ng) expression vectors. For analysis of the synergistic effect of Cdx1 and Zic2, N2a cells were transfected with 100 ng of wild type or mutated *Pax3*NCE2-Luciferase reporter construct alone, or with fixed amount of $_{FLAG}$Cdx1 (5ng) and/or $_{HA}$Zic2 (10 ng) expression vectors. When required, an empty expression vector was also included to complete the final amount of DNA per well to 150 ng. For analysis of the synergistic transactivation of *Pax3*NCE2 by Cdx1, Zic2 and Sox2 co-expression, fixed amounts of $_{FLAG}$Cdx1 (10 ng), $_{HA}$Zic2 (40 ng) and $_{Myc}$Sox2 (50 ng) were used and an empty expression vector was also included to ensure a total of 200 ng of DNA per well. All transfections were performed at least five times in triplicate. In all experiments, specificity of NCE2 activation was assessed by using a previously described luciferase reporter construct consisting of the *Pax3* 150bp minimal promoter alone (Sanchez-Ferras et al., 2012), as a negative control. Forty-eight hours after transfection, cells were disrupted in 100 µl of lysis buffer (0.1 M Tris [pH 8.0], 1% Igepal, 1 mM dithiothreitol) and assessed for luciferase activity with a Berthold LB9507 luminometer (Berthold Technologies).

Expression of exogenous $_{FLAG}$Cdx1, $_{HA}$Zic2 and $_{Myc}$Sox2 in cell lysates was verified by western blot.

To modulate Zic2 activity, N2a cells were seeded in 100-mm tissue culture plates ($2x10^6$ cells per plate) and transiently transfected with increasing doses (3, 6 and 9 μg) of $_{HA}$Zic2-IRES-GFP expression vector. Forty-eight hours after transfection, GFP-positive cells were recovered by FACS and analyzed by RT-PCR (for *Pax3* and *Gapdh* expression) as well as via western blot (for $_{HA}$Zic2 and Gapdh expression).

For dose-response analyses of Shh-induced repression of *Pax3*, *Zic2* and *Zic5* expression, N2a cells were seeded in 6 well plates ($4x10^5$ cells per well) and treated the following day with varying doses (50-200 ng/ml) of recombinant mouse Shh-N proteins (R&D systems) in absence or presence of 1ug/ml cyclopamine (R&D systems). Cells were harvested 24h post-treatment and analyzed by RT-PCR.

For cycloheximide (CHX) treatments, N2a cells were pre-treated for 30 min with 30 μg/mL of CHX or with the vehicle (DMSO) alone as previously described (Sanchez-Ferras et al., 2012) and then treated for 24h with or without Shh-N recombinant protein (200 ng/mL) in the presence of 1 μg/mL of CHX.

For rescue experiments, N2a cells were seeded in 100-mm tissue culture plates ($2x10^6$ cells per plate) and transiently transfected with $_{HA}$Zic2-IRES-GFP or empty IRES-GFP expression vectors. The following day, cells were incubated in the presence of Ctl or Shh-enriched (200 ng/ml) medium and cultured for another 24h prior to recovery of GFP-positive cells by FACS.

RNA isolation and RT-PCR analyses were performed as previously described (Sanchez-Ferras et al., 2012). The primers used in this study were: *Pax3* (forward: 5'-CCTGCCAACATACCAGCTGTCG-3', reverse: 5'-CTGAGGTGAAAGGCCATTGCCG-3'); *Zic2* (forward:5'-AAGGTCTTCGCACGCTCCGAG-3', reverse: 5'-

CGCAACGAGCTGGGATGCGTGT-3'); *Zic5* (forward: 5'-CAAGATCCACAAGCGCACTCATACA-3', reverse: 5'-TTGGGTCCAGCACAGGGGACAAAG-3'); *Gapdh* (forward: 5'-TCCTGCACCACCAACTGCTTAGC-3', reverse: 5'-AGGTCCACCACCCTGTTGCTGTA-3').

2.3.8 Statistical analyses

Statistical analyses were performed using GraphPad Prism software version 5.0. Differences between means were evaluated by one way ANOVA followed by a Tukey's post test and classed as not significant (n.s., $p>0.05$) or significant (*, $p<0.05$; **$p<0.01$; ***$p<0.001$).

2.4 Results

2.4.1 A positive Cdx input is necessary but not sufficient to refine the Pax3 expression domain in the caudal neuroectoderm

Cdx genes (*Cdx1/2/4*) encode homeodomain transcription factors exhibiting functional overlap and overlapping expression patterns in the caudal embryo (Beck et al., 1995; Gamer and Wright, 1993; Meyer and Gruss, 1993; Savory et al., 2011a; Savory et al., 2009b; van den Akker et al., 2002; van Nes et al., 2006). Using mouse transgenic reporters and murine cell line perturbation analyses, we have previously demonstrated that Cdx proteins are required, downstream of canonical Wnt signaling, to induce *Pax3* expression in the mouse caudal neuroectoderm (Sanchez-Ferras et al., 2012). However, as shown in Figure 2.1A by immunofluorescence analysis of Cdx2 and Pax3 distribution in e9.5 caudal NT sections, *Cdx* and *Pax3* expression patterns only partially overlap (Beck et al., 1995; Gamer and Wright, 1993; Goulding et al., 1991; Meyer and Gruss, 1993). Indeed, Cdx proteins are broadly distributed in the entire tailbud and, more anteriorly, they are detected along the whole DV axis of the recently closed NT. In contrast, *Pax3* expression is restricted to the lateral borders of the open PNP and, more anteriorly, to the dorsal NT. This implies that non-Cdx

inputs are required to refine the *Pax3* expression domain in the caudal neuroectoderm and leaves us with three possibilities that may involve either one or two mechanisms (Figure 2.1B). The first possibility (Mechanism-1 only) would involve one or several Cdx cofactors that exhibit more-or-less restricted neural expression in the lateral PNP and dorsal NT. The second possibility (Mechanism-2 only) would involve a negative input from a repressor expressed in the medial PNP and ventral NT. The third possibility would involve a combination of both mechanisms.

2.4.2 The neural and dorsally restricted zinc finger transcription factor Zic2 regulates Pax3 expression

Based on their overlapping expression pattern with Cdx and Pax3, we first hypothesized that Zic2/5 transcription factors might contribute to Mechanism-1. Indeed, Zic proteins can transactivate target genes and of the five murine *Zic* genes, only the *Zic2-Zic5* gene pair (sharing the same 5'-flanking region) is expressed in the caudal most region of the neuroectoderm (Merzdorf, 2007). Moreover, *Pax3* expression has been previously reported to be significantly reduced in NCC homozygous for a knockdown allele of *Zic2* while *Zic2*-null mutants exhibit posterior NT and NCC defects that phenocopy those observed in *Pax3*-null mutants (Nagai et al., 2000). To validate the candidacy of Zic2 as a *Pax3* regulator, we began by carrying out a detailed analysis of their expression pattern using whole-mount *in situ* hybridization in stage-matched e9.0 and e9.5 mouse embryos. In accordance with previous studies, this analysis revealed a striking and extensive overlap in the neural expression domain of both genes along the AP axis (Figure 2.2 A-J). Interestingly, as evidenced by its more posterior limit of expression in the open PNP, this analysis notably demonstrated that *Zic2* is in fact induced earlier than *Pax3* in the caudal embryo (compare Figures 2.2C,E,G with 2.2 D,F,H). Following NT closure, *Pax3* and *Zic2* are then found to be equally expressed in a dorsal domain that covers ~40% of the DV neural axis (bracket in Figure 2.2I-J). It is noteworthy that such overlapping expression pattern is neural specific since, outside the NT, *Pax3* and *Zic2*

are expressed in a mutually exclusive manner in the somites (asterisk in Figure 2.2I,J). Thus, these data support the candidacy of Zic2 and further suggest that it might be involved in both the induction and dorsal restriction of neural *Pax3* expression. To more formally test our hypothesis that Zic2 regulates *Pax3* expression, we then overexpressed Zic2 in the murine NCC-derived N2a cell line and evaluated the effect on *Pax3* expression by RT-PCR. As shown in Fig.2K, overexpression of Zic2 resulted in a specific and dose-dependent increase of *Pax3* expression. In summary, our data strongly suggest that Zic2 is a regulator of *Pax3* expression.

2.4.3 Zic2 directly binds and transactivates Pax3NCE2

We previously showed that the *Pax3*NCE2 CRM is sufficient for recapitulating induction and dorsal restriction of *Pax3* expression in the caudal neuroectoderm (Sanchez-Ferras et al., 2012). As we hypothesised that Zic2 is a Cdx neural cofactor, we reasoned that Zic2 should be able to activate *Pax3* expression via this Cdx-dependent CRM. We tested this possibility via luciferase assays in N2a cells and found that Zic2 transactivated a *Pax3*NCE2 reporter in a dose-dependent manner, similarly to endogenous *Pax3* (Figure 2.3A and 2K). EMSAs were then used to determine whether Zic2 directly binds to *Pax3*NCE2. Zic proteins are known to bind GC-rich elements but very few ZicBS have been previously identified in CRMs of Zic-regulated genes (Merzdorf, 2007). Thus, to avoid excluding any unpredicted binding site, we scanned the whole NCE2 sequence with eight previously described overlapping probes (Sanchez-Ferras et al., 2012). This approach revealed that one of these probes was bound by Zic2 as efficiently as a probe containing a previously described ZicBS (Mizugishi et al., 2001) (Figure 2.3B). We confirmed the presence of $_{HA}$Zic2 in the complex formed with the NCE2 probe by supershift with an anti-HA antibody (Figure 2.4E'). In accordance with previous reports, sequence analysis of the NCE2 probe revealed the presence of a GC-rich 9-nucleotide core sequence (5'-CTGCTGGGG-3'). Mutation of guanosine residues in position 6 and 7 prevented Zic2 binding to a cold probe bearing these point mutations, thus demonstrating

specificity of Zic2 binding to this NCE2 target sequence (Figure 2.4E'). We next sought to determine whether Zic2, like Cdx proteins, occupies the endogenous *Pax3*NCE2 using ChIP-PCR. To this end – and because we failed to obtain a ChIP-grade anti-Zic2 antibody – we overexpressed $_{HA}$Zic2 in N2a cells and evaluated its presence on NCE2 using anti-HA. $_{FLAG}$Cdx1-overexpressing cells were used for comparison purposes. As shown in Fig.3C, *Pax3*NCE2 was specifically amplified in sonicated chromatin samples immunoprecipitated with either anti-HA or anti-FLAG antibodies, while no amplification was obtained when a mouse pre-immune serum was used for IP. Altogether, these results demonstrate that Zic2 directly binds and transactivates *Pax3*NCE2.

2.4.4 Zic2 directly interacts with Cdx1

Analysis of NCE2 sequences revealed that the ZicBS is flanked by CdxBS in close proximity (Figure 2.5B), further supporting a putative Cdx cofactor function for Zic2. To directly test this possibility, we carried out protein-protein interaction assays beginning with co-IP in $_{HA}$Zic2- and/or $_{FLAG}$Cdx1-transfected Cos7 cells. As shown in Figure 2.4A, mouse Zic2 was found to specifically interact with Cdx1 when both are co-expressed. Then, to assess whether this interaction was direct, we performed GST-based pull-down assays using *in vitro* translated FLAG-tagged human ZIC2 and bacterially produced GST fusion proteins containing either the full-length Cdx1, the Cdx1 homeodomain or the Cdx1 N-terminal domain (GST-Cdx1, GST-Cdx1Homeo and GST-Cdx1Nterm). Western blotting using an anti-FLAG antibody showed that $_{FLAG}$ZIC2 directly interacts with both the full-length Cdx1 and the Cdx1 homeodomain whereas no interaction was observed with the Cdx1 N-terminus (Figure 2.4B). Mapping of Cdx1 and ZIC2 interacting domains was next assessed using a series of *in vitro* translated $_{FLAG}$ZIC2 deletion-containing proteins and the GST-Cdx1Homeo fusion protein (Figure 2.8A). This analysis revealed that ZIC2 and Cdx1 interact via their respective DNA binding domain since no interaction with the Cdx1 homeodomain was observed when the Zinc finger domain of ZIC2 was deleted

(construct 1-255) (Figure 2.4C-D and Figure 2.8B). Regions located N-terminus (constructs 140-532 and 255-532) or C-terminus (1-415) to the Zinc finger domain have apparently no major impact on this interaction (Figure 2.4C-D). Intrigued by this discovery, we then tested whether Zic2 and Cdx1 could still physically interact with each other in the presence of their respective target sequences using EMSA as well as GST pull-down assays. In EMSA, we found that pre-incubation of $_{HA}$Zic2-with $_{FLAG}$Cdx1-expressing Cos7 nuclear extracts did not preclude target DNA binding and even resulted in the appearance of novel bands of higher molecular weight either on the ZicBS (E'), the CdxBS1 (E'') or the CdxBS3 (E'''), suggesting the formation of a Zic2-Cdx1-DNA complex on these elements. Consistent with these outcomes, pre-incubation of $_{FLAG}$ZIC2 with increasing amounts of the whole NCE2 did not impair its interaction with the Cdx1 homeodomain in GST pull-down assays (Figure 2.8). These data demonstrate that, although the Zic2-Cdx1 physical interaction is mediated by their respective DNA binding domain, this does not prevent the ability of both Zic2 and Cdx1 to bind their respective target sequences.

2.4.5 Robust Zic2 and Cdx1 functional interaction on Pax3NCE2 requires a positive input from the neural specific Sox2 transcription factor

Last validation of Zic2 as a Cdx1 cofactor required the demonstration of a functional interaction with Cdx1 and we assessed this via luciferase assays in both N2a and undifferentiated P19 cell lines. In N2a cells, Cdx1-Zic2 co-expression resulted in a weak synergistic activation of the *Pax3*NCE2 luciferase reporter (Figure 2.5A). Interestingly, mutation of the three functional CdxBS alone or in combination with mutation of the ZicBS significantly abrogated this Cdx1-Zic2 cooperative effect. This is consistent with the central role of Cdx1 in activation of *Pax3*NCE2 (Sanchez-Ferras et al., 2012) and suggests that binding of Cdx1 to this CRM is important for the Cdx1-Zic2 functional interaction. Intriguingly, although mutation of the ZicBS also abrogated the Cdx1-Zic2 cooperative effect, these analyses further revealed that Zic2 ability to transactivate *Pax3*NCE2 can be independent of the ZicBS as well as

the three CdxBS (Figure 2.5A). Moreover, in the undifferentiated P19 cell line, Zic2 failed to transactivate *Pax3*NCE2 and no cooperative effect between Cdx1 and Zic2 was observed (data not shown). Therefore, this suggests that interaction with another neural factor is necessary for the Zic2-Cdx1 complex to robustly transactivate *Pax3*NCE2. Based on several lines of evidence, Sox2 quickly emerged as a primary candidate in this regard. Indeed, at least one potential SoxBS is present in *Pax3*NCE2 (Figure 2.5B), SoxB1 family members have recently been implicated as positive regulators of *Pax3* expression (Moore et al., 2013) and the SoxB1 family member Sox2 has been shown to directly interact with Cdx1 (Beland et al., 2004). To assess the putative contribution of Sox2, we first evaluated whether this transcription factor alone was able to transactivate the *Pax3*NCE2 luciferase reporter in N2a cells. Consistent with our hypothesis, Sox2 strongly transactivated *Pax3*NCE2 in a dose-dependent manner (Figure 2.5C). Of note, such effect was not observed in P19 cells, suggesting that the Sox2 capacity to transactivate this CRM relies on the presence of other, as yet undefined, neural factors (Figure 2.5C). To demonstrate the importance of the Sox2 input in the functional interaction between Cdx1 and Zic2, we co-transfected N2a cells with Cdx1, Zic2 and Sox2 and evaluated the activation of the *Pax3*NCE2 luciferase reporter. Interestingly, a significant synergistic effect was observed when the three factors were co-expressed (Figure 2.5D; see Figure 2.9 for western blot validations). Together, these results suggest that *Pax3*NCE2 is robustly activated by a concerted interaction between Cdx1, Zic2, and Sox2 trans-acting factors.

2.4.6 Zic2 is a potential mediator of the Shh-induced repression of Pax3

Having confirmed that Zic2 is a *bona fide* Cdx cofactor, and thus contribute to Mechanism-1, we next investigated the possibility that it could also contribute to the repressor-dependent Mechanism-2 (Figure 2.1B). Indeed, a ventrally Shh-induced repressive pathway is a prominent candidate to consider and both *Zic* members and *Pax3* are known to be affected by such regulatory cascade. Studies in multiple

species, including mice, indicate that Shh signals secreted by the notochord and floorplate are required to prevent ectopic expression of *Pax3* and *Zic* members in the ventral NT (Aruga et al., 2002b; Brown et al., 2003; Goodrich et al., 1997; Goulding et al., 1993; Holtz et al., 2013; Moore et al., 2013; Rohr et al., 1999). While the molecular mechanism of this regulation is poorly understood for *Zic* members, recent work in zebrafish and chick embryos suggest that the Shh target Nkx6.1 represses *Pax3* expression via direct binding to a *Pax3* CRM located in the 4th intron (Degenhardt et al., 2010; Moore et al., 2013). However, whether such Shh-Nkx6.1 circuit is at work in the mouse and active at all AP levels has not been investigated. To verify this, we compared the expression profile of *Pax3* and *Zic2* with the distribution of Nkx6.1 proteins in the caudal neuroectoderm of stage matched e9.5 mouse embryos (Figure 2.6A-O). Surprisingly, we found that Nkx6.1 is not detected in the open PNP and recently closed NT whereas *Pax3* is already induced and properly restricted in these structures (compare Figure 2.6A-B with K-L). Therefore, the Shh-Nkx6.1 pathway cannot be responsible for the restriction of *Pax3* expression in the caudal neuroectoderm of mouse embryos. On the other hand, we found that *Zic2* is dynamically expressed along the D-V axis of the caudal neuroectoderm. *Zic2* transcripts are indeed first detected in the whole PNP and along the entire DV axis of the recently closed NT before becoming dorsally restricted more anteriorly (Figure 2.6F-J). In more anterior regions, a perfect overlap is finally observed between *Zic2* and *Pax3* neural expression domains (Figure 2.2I-J). Very interestingly, we also noticed that the onset of *Zic2* repression in the ventral NT is correlated with the onset of Nkx6.1 expression (see arrow in Figure 2.6H and M). On one hand, all these observations suggest that, while the Cdx-Zic2 neural input is necessary for *Pax3* induction, a Shh-Nkx6.1-independent input must also be required to restrict *Pax3* expression in the open PNP and most caudal NT. On the other hand, these expression data also strongly suggest that Zic2 might be an intermediary in the mechanism of *Pax3* repression by the Shh-Nkx6.1 pathway. We sought to validate this hypothesis

using the murine N2a cell line and first validated this line as a good model to study the mechanism of *Pax3* repression by Shh (Figure 2.10). We then performed a dose-response assay to evaluate whether *Zic2* and *Pax3* were both repressed by increasing doses of Shh (Figure 2.6P). As expected from previous work (Aruga et al., 2002b; Brown et al., 2003; Goodrich et al., 1997; Goulding et al., 1993; Holtz et al., 2013; Moore et al., 2013; Rohr et al., 1999) , Shh repressed the expression of *Pax3* as well as expression of the *Zic2-Zic5* gene pair. To finally confirm our hypothesis, we tried to rescue *Pax3* expression by overexpressing Zic2 in Shh-treated cells. In agreement with a Zic-dependent mechanism, transfection with a Zic2 expression vector specifically rescued *Pax3* expression from the repressive input of Shh (Figure 2.6Q). Taken together, this evidence strongly support an intermediary role for Zic2 in the mechanism of *Pax3* repression by Shh in the mouse.

2.5 Discussion

In this work, we report a novel direct role for Zic2 in the regulation of *Pax3* neural expression. Our data indicate that Zic2 can transactivate *Pax3* at least via the NCE2 CRM, which we previously showed to be sufficient for recapitulating caudal *Pax3* expression in pre-migratory NCC and dorsal NT (Sanchez-Ferras et al., 2012). Interestingly, we also provide evidence that Zic2 is involved in both the Wnt-mediated induction (by interacting with Cdx proteins) and the Shh-induced dorsal restriction (by being itself repressed by Shh signals) of *Pax3* expression. Therefore, our studies provide important mechanistic insights into how AP and DV instructive cues can be coordinately integrated on a short CRM.

2.5.1 Role of the Zic2-Cdx complex in Wnt-mediated induction of Pax3 neural expression

In previous studies, we presented evidence that a Wnt-Cdx circuit is critically required for activation of a *Pax3*NCE2 reporter in transgenic mice (Sanchez-Ferras et al., 2012). These data support a model in which Cdx proteins convey the

posteriorizing Wnt-ßCatenin signals to *Pax3* in order to mediate induction of trunk NCC (Pilon et al., 2006; Pilon et al., 2007; Sanchez-Ferras et al., 2012). Interestingly, indirect activation of *Pax3* expression by canonical Wnt signaling – involving either AP2a or Gbx2 as intermediates – has also been described in *Xenopus* embryos during induction of cranial NCC (de Croze et al., 2011; Li et al., 2009). On the other hand, other groups have in addition reported that *Pax3* could be directly regulated by Wnt-ßCatenin signaling through CRMs distinct from NCE2 (Degenhardt et al., 2010; Moore et al., 2013; Zhao et al., 2013). This suggests that regulation of *Pax3* expression by the canonical Wnt pathway could be controlled via both direct and indirect means in order to buffer against perturbations in the Wnt input. In the caudal neuroectoderm, a third level of regulation would even be possible given that Cdx proteins can directly and positively interact with the ßCatenin-Lef/Tcf complex on Lef/Tcf binding sites (Beland et al., 2004). However, the very broad distribution of Cdx proteins in both neural and non-neural cells of the caudal embryo (Meyer and Gruss, 1993; Takada et al., 1994) implies that other players must cooperate with the Wnt-Cdx inductive input to spatially refine the *Pax3* expression domain.

The present study demonstrates that Zic2 not only acts as a *Pax3* regulator, but also as a neural-specific Cdx cofactor. Indeed, we found that Zic2 and Cdx1 cooperate in the activation of *Pax3*NCE2 and physically interact via their respective DNA binding domain (Figure 2.4 and 2.5). Such involvement of DNA binding domains in protein-protein interactions is not unusual and examples implicating either the Zic2 C2H2 zinc finger region or the Cdx1 homeodomain have been previously described (Beland et al., 2004; Pourebrahim et al., 2011). On the other hand, to the best of our knowledge, this is the first report showing that the Zic zinc finger region or the Cdx homeodomain can simultaneously mediate protein-protein as well as protein-DNA interactions (Figure 2.4, 2.5 and Figure 2.8). This bifunctionality is most likely due to the fact that, for each of these proteins, establishment of direct contacts with DNA bases is not carried out by the whole DNA binding domain *per se* but by specific

subregions only. Indeed, previous studies have suggested that only zinc fingers 3-5 of Zic2 might be implicated in DNA recognition (Houtmeyers et al., 2013; Merzdorf, 2007; Mizugishi et al., 2004) whereas, for homeodomain proteins, DNA binding activity is known to be specifically conferred by the 3rd α-helix (Gehring et al., 1994). Thus, this suggests that the Zic2-Cdx1 interaction might be mediated by the zinc fingers 1-2 of Zic2 and the 1st and/or 2nd α-helix of the Cdx1 homeodomain. Regardless of the exact minimal interacting domain on each protein, it should be noted that functional interactions between different members of each family are expected given the very high conservation of both Zic and Cdx DNA binding domains (Houtmeyers et al., 2013; Lohnes, 2003).

Interestingly, our data also indicate that, although both Zic2 and Cdx1 have the ability to bind *Pax3*NCE2 sequences in a direct as well as in an indirect manner (Figure 2.4), robust activation by the Zic2-Cdx1 complex requires direct binding of each protein to this CRM as well as the participation of Sox2 (Figure 2.5) – a previously described Cdx interacting partner (Beland et al., 2004). Such contribution of Sox2 is in line with recent work suggesting the implication of SoxB members in the regulation of neural *Pax3* expression via another CRM located in the 4th intron (Moore et al., 2013). On the other hand, it is interesting to note that in contrast to Cdx proteins (Sanchez-Ferras et al., 2012), neither Zic2 nor Sox2 can activate *Pax3* expression in non-neural cells (Figure 2.5C and data not shown). Taken together with the indispensable role of CdxBS1-3 for *Pax3*NCE2 activity (Sanchez-Ferras et al., 2012), this strongly suggests that Cdx proteins have a pioneer-like activity required to attract and tether neural-specific transcriptional activators on this CRM (Zaret and Carroll, 2011). Interestingly, a similar pioneer-like activity has been also previously suggested for Cdx2 in the regulation of the transcriptional program of intestinal epithelial cells (Verzi et al., 2013) as well as for Pdx1 – another member of the Cdx-containing paraHox family – in the regulation of several genes important for

pancreatic ß cell development and function (Hoffman et al., 2010; Ohneda et al., 2000).

2.5.2 Restriction of Pax3 expression to the lateral borders of the PNP and the dorsal NT

A role for notochord- and floor plate-secreted Shh in the regulation of *Pax3* restriction in the dorsal NT has been described in diverse model organisms (Goulding et al., 1993; Litingtung and Chiang, 2000a; Moore et al., 2013). In the mouse, Shh gain-of-function in *Ptch1*-null mutants results in a nearly complete loss of *Pax3* expression whereas Shh loss-of-function in Shh-null embryos results in expansion of *Pax3* expression along the entire DV neural axis (Chiang et al., 1996; Goodrich et al., 1997). Recent studies suggest that this Shh-induced repressive input could be mediated by the Nkx6.1 homeodomain transcriptional repressor – a known direct target of the Shh-Gli pathway (Peterson et al., 2012). Indeed, Moore et al. (Moore et al., 2013) reported that Nkx6.1 can directly bind a *Pax3* CRM located in intron-4 (called CNE3) and that mutation of a putative binding site for homeodomain proteins within this CRM is required to turn on ubiquitous but mosaic expression of a reporter gene driven by this CRM in zebrafish embryos. Combination of this work with Tcf3 ChIP-seq data from ES cells (Marson et al., 2008) led these authors to propose a model whereby Shh-induced Nkx6.1 could counteract canonical Wnt signaling in order to restrict induction of *Pax3* expression to the lateral borders of the PNP (Moore et al., 2013). However, in addition to the fact that Tcf3 is not a direct mediator of Wnt-ßCatenin positive inputs (Wu et al., 2012), this model is not consistent with the Nkx6.1 expression profile in the mouse. Indeed, we and others have shown that Nkx6.1 is not expressed in the open PNP and recently closed NT, whereas *Pax3* is already induced and properly restricted in these regions (Figure 2.6) (Peterson et al., 2012). Moreover, *in situ* hybridization analyses have failed to detect *Shh* transcripts in the caudal neuroectoderm (Gofflot et al., 1997), implying that no other Shh targets could be involved in this mechanism. In agreement with this, *Zic2*, a

known Shh-regulated gene in the mouse (Brown et al., 2003), is broadly expressed in the entire PNP and recently closed NT (Figure2.6). Taking into account all these observations, the proposed Shh-Nkx6.1 repressive circuit cannot contribute to the induction phase of *Pax3* expression and can only be implicated in the maintenance of *Pax3* dorsal restriction.

Interestingly, our data indicate that an alternative pathway involving repression of the *Pax3* activator Zic2 could also mediate the Shh-induced repression of *Pax3* expression (Figure 2.6). This hypothesis is strongly supported by data from Moore et al. (Moore et al., 2013) showing that Nkx6.1 overexpression has a much more dramatic impact on repression of *Pax3* expression in comparison to the CNE3 reporter bearing the mutated homeodomain binding site. Moreover, our expression data in Figure 2.6 – showing that the onset of *Zic2* repression is correlated with the onset of Nkx6.1 expression in the ventral NT – are also consistent with the idea that *Zic2* might be a direct target of Nkx6.1. Although it is currently highly speculative, we propose that this alternative Shh-Nkx6.1-Zic2-Pax3 pathway targeting NCE2 would be functionally redundant with the more straightforward Shh-Nkx6.1-Pax3 pathway targeting CNE3 (Moore et al., 2013). Given the co-expression of the *Zic2-Zic5* gene pair as well as the known functional overlap between Zic members (Aruga et al., 2002a; Inoue et al., 2004; Inoue et al., 2008; Inoue et al., 2007a; Inoue et al., 2007b), we also believe that Zic5 acts redundantly with Zic2 in this mechanism. Furthermore, given that Zic and Gli proteins have similar DNA binding specificity and Zic inhibit Gli transcriptional activity (Mizugishi et al., 2001), it is tempting to speculate that cross-repression between Zic2/5 and the Shh pathway would be important for proper D-V patterning of the NT.

2.5.3 Coordinated integration of A-P and D-V instructive cues and redundancy between Pax3 CRMs

As evidenced by our work and the work of others, *Pax3* gene expression is regulated by two sets of CRMs that are clustered in the 5'-flanking region or in intron-4, respectively (Figure 2.5B) (Degenhardt et al., 2010; Garnett et al., 2012; Milewski et al., 2004; Moore et al., 2013; Natoli et al., 1997; Sanchez-Ferras et al., 2012; Zhao et al., 2013). Interestingly, although the requirement of intron-4 CRMs has not yet been formally addressed via gene targeting approaches, deletion of 5'-located CRMs in the mouse is indicative of functional overlap between each regulatory cluster (Degenhardt et al., 2010). It is also interesting to note that, although each set of CRMs is expected to respond to the same A-P and D-V instructive cues, different mechanisms appear to be used (e.g. Shh-Nkx6.1-Pax3 vs Shh-Nkx6.1-Zic2-Pax3) and different complementarities are seen among CRMs of each cluster. Within the 5' cluster, NCE1 and NCE2 appear to be complementary for the spatial control of *Pax3* expression along the A-P axis. Indeed, while we have shown that NCE2 drives reporter gene expression in a caudal-specific and Cdx-dependent manner, others have shown that NCE1 is involved in the regulation of cranial *Pax3* expression in a Pbx1-dependent manner (Chang et al., 2008; Sanchez-Ferras et al., 2012). In contrast, within the intron-4 cluster, CNE3 (also known as ECR2 or IR1) and CNE1 (also known as IR2) appear to exhibit temporal complementarities in the control of *Pax3* expression, with CNE3 suggested to be involved in the induction phase in a Lef/Tcf-dependent manner and CNE1 suggested to be involved in the maintenance phase in a Pax3/7-dependent manner (Degenhardt et al., 2010; Garnett et al., 2012; Moore et al., 2013). In accordance with the key developmental role of Pax3, such functional overlap at both the *cis* and *trans* level most likely allows to protect its expression from the otherwise deleterious impacts of genetic mutations in *cis* or perturbations of signaling pathways in *trans*.

Our work on NCE2 indicates that it may serve as a good model to further understand how AP and DV instructive cues can be coordinately integrated on a single CRM. Indeed, in addition to allow recapitulating induction of *Pax3* expression in the PNP, NCE2-driven expression is properly restricted (Sanchez-Ferras et al., 2012). It is also noteworthy that we and others have shown that NCE2 seems particularly enriched in binding sites for a wide array of transcription factors including members of the Cdx, Pou class-III, SoxB, Zic and Tead families (Milewski et al., 2004; Pruitt et al., 2004; Sanchez-Ferras et al., 2012). In an effort to gather all this information into a dynamic network (Figure 2.7), we propose a model in which the posteriorizing cue from the Wnt-Cdx circuit, in cooperation with Zic and SoxB positive inputs, is first required to induce *Pax3* expression in the PNP. Later on (i.e. more rostrally), the Shh-Nkx repressive input is then most likely involved in the maintenance of the dorsally-restricted *Pax3* expression domain via repression of *Zic2/5* expression. In absence of Cdx factors more rostrally, maintenance is also most likely ensured by the positive input of other transcription factors such as members of the Pou class III and Tead families (Milewski et al., 2004; Pruitt et al., 2004). While this model clearly allows understanding how the *Pax3*NCE2 CRM is regulated by A-P instructive cues (Wnt-Cdx), this is much less clear regarding the integration of DV instructive cues. Indeed, our studies clearly show that an unknown Shh-independent factor is required during the induction phase in order to limit expression to the lateral borders of the PNP. More work will definitely be necessary for the identification of this key player, once again in accordance with the mechanisms described in Figure 2.1B.

2.6 Conclusion

This study allowed confirming the central role of Cdx proteins in the induction of *Pax3* neural expression as well as identifying novel molecular functions for Zic2 as a Cdx co-factor and activator of *Pax3* expression. By allowing extension of the list of transcription factors able to bind and transactivate *Pax3*NCE2, our work also suggests that this CRM behaves as a "super-enhancer" region.

2.7 Acknowledgements

The authors thank Denis Flipo (UQAM) for FACS analyses and assistance with confocal imaging. Jonathan Epstein (Penn Cardiovascular Institute), Sabine Tejpar (Katholieke Universiteit Leuven), Jun Aruga (RIKEN Brain Science Institute), Mahmud Bani-Yaghoub (University of Ottawa) and David Lohnes (University of Ottawa) are thanked for kindly agreeing to provide expression vectors. This work was supported by a grant from the Canadian Institute for Health Research (CIHR grant number MOP-111130) to NP. OSF holds an Alexander-Graham-Bell scholarship from the Natural Science and Engineering Research Council (NSERC) of Canada. NP is a Fonds de la Recherche du Québec – Santé (FRQS) Junior2 scholar.

2.8 Figures

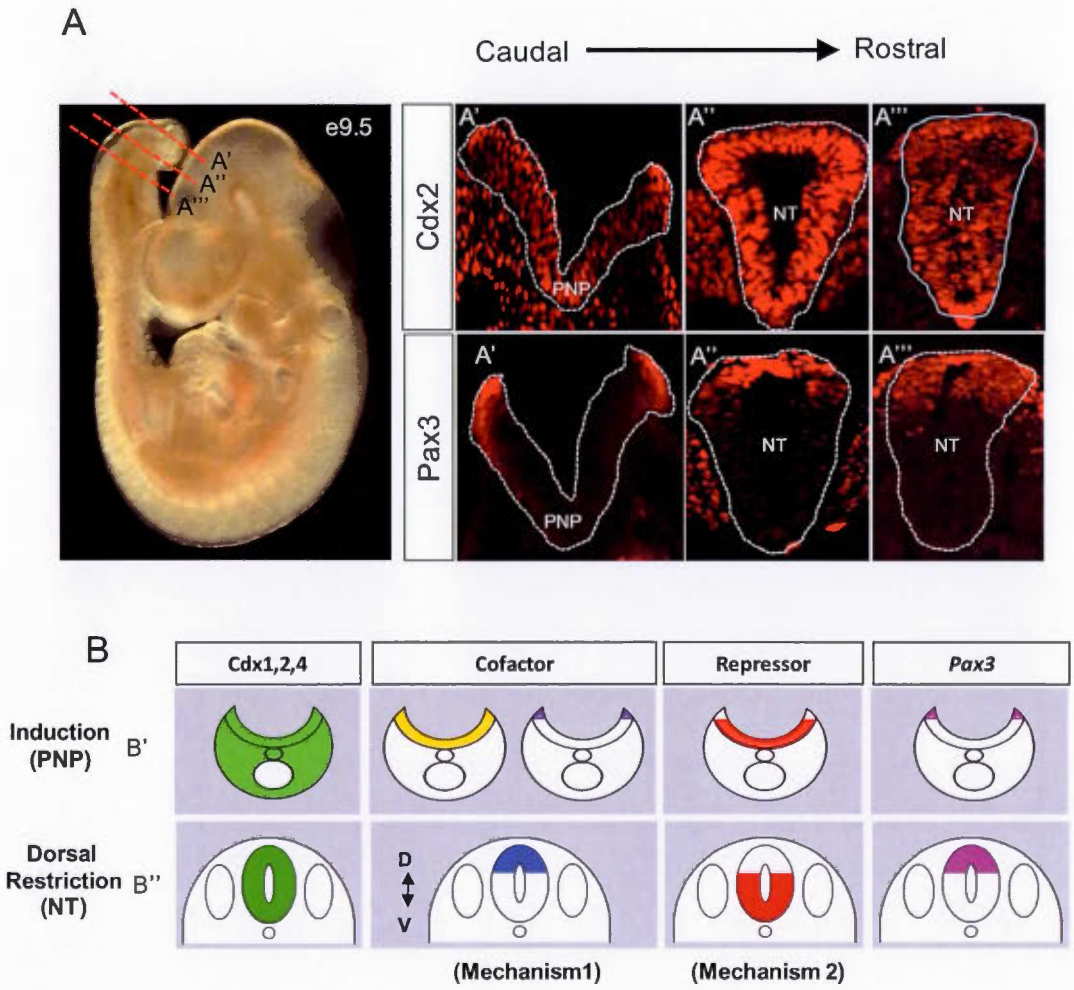


Figure 2.1 Partially overlapping Cdx and Pax3 expression domains indicate that Cdx proteins are necessary but not sufficient to establish the *Pax3* expression domain in the caudal neuroectoderm.

(A) Immunofluorescence analysis of Cdx2 and Pax3 distribution in tailbud sections of a 25-somite stage mouse embryo (e9.5). The dashed red lines in the left panel indicate the AP level at which the sections shown on the right were made *(A', A'', A''')*. The dotted white lines in the right panel delimit the posterior neural plate (PNP) and the neural tube (NT) structures. Note that Cdx2 is broadly expressed in the tailbud including the whole PNP (A') and, more rostrally, in the entire NT (A'', A''') while Pax3 domain is limited to the lateral borders of the PNP (A') and to the dorsal part of the NT (A'', A'''). *(B)* Schematic drawing of two putative mechanisms to explain the induction and dorsal restriction of *Pax3* neural

expression (pink) via Cdx proteins (green). *(B') Pax3* expression at the lateral borders of the PNP might be established by collaboration between Cdx proteins and a positive input from a neural-specific protein with more-or-less restricted expression: i.e. a Cdx cofactor expressed in the PNP (yellow) or at the lateral borders of the PNP (purple) [Mechanism 1]. Restriction of *Pax3* expression at the PNP border may also be achieved by integration of the positive Cdx input with the negative input from a repressor expressed in the medial PNP (red) [Mechanism 2]. *(B'')* At more rostral levels, where Cdx expression is restricted to the NT, the Cdx input might be integrated with the positive input from a dorsally restricted Cdx cofactor (blue) [Mechanism 1] and/or the negative input from a ventrally expressed repressor (red) [Mechanism 2].

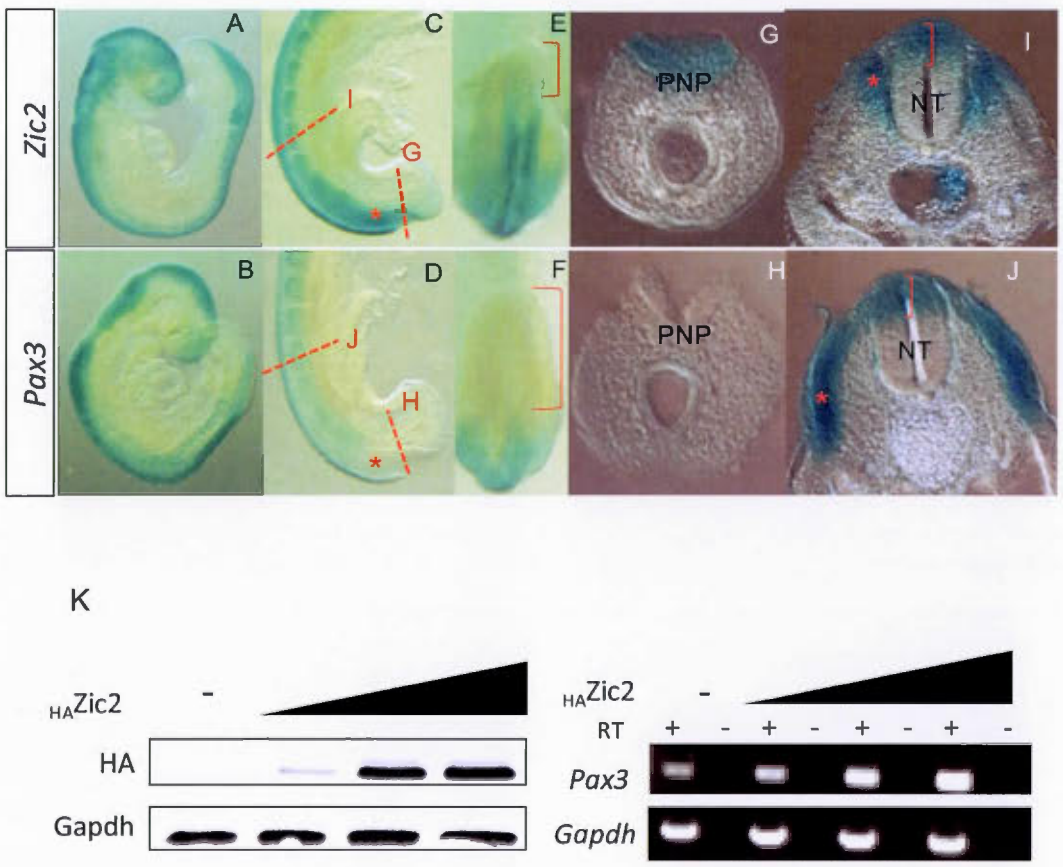


Figure 2.2 The Zic2 transcription factor regulates *Pax3* expression.

(A-J) Overlapping spatiotemporal *Zic2* and *Pax3* expression patterns are consistent with a role for Zic2 in regulating *Pax3* expression. *(A-B)* Lateral view of 20-somite stage embryos (~e9.0) with anterior to the left showing *Zic2 (A)* and *Pax3 (B)* expression domains assessed by whole-mount *in situ* hybridisation. *(C-J)* Comparison of *Zic2* and *Pax3* expression domains in the caudal end of 24-somite stage embryos (~e9.5). The dotted lines in the lateral views shown in C-D indicate the level at which the transverse sections shown in G-J were cut. The red brackets in E-F designate the length of the caudal tip of the tailbud region (dorsal view) that is devoid of *Zic2* and *Pax3* gene expression. Note that, compared to *Pax3*, the *Zic2* expression domain in the open PNP extends more posteriorly (C-H). Also note the extensive overlap in *Pax3* and *Zic2* expression in the dorsal NT (C-D; red brackets in I-J). *Pax3* and *Zic2* expression patterns do not overlap in the developing somites. The presomitic mesoderm (C, D) and somites (I, J) are denoted by asterisks. *(K)* Analyses of endogenous *Pax3* expression in N2a cells transfected with increasing amounts of a $_{HA}$Zic2-IRES-GFP expression vector. GFP-positive N2a cells were recovered by FACS and exogenous $_{HA}$Zic2 as

well as endogenous *Pax3* expression were analyzed by western blot and RT-PCR, respectively. Note that increasing levels of $_{HA}$Zic2 (left panel) result in a dose-dependent increase of endogenous *Pax3* expression but not *Gapdh* used as control (right panel).

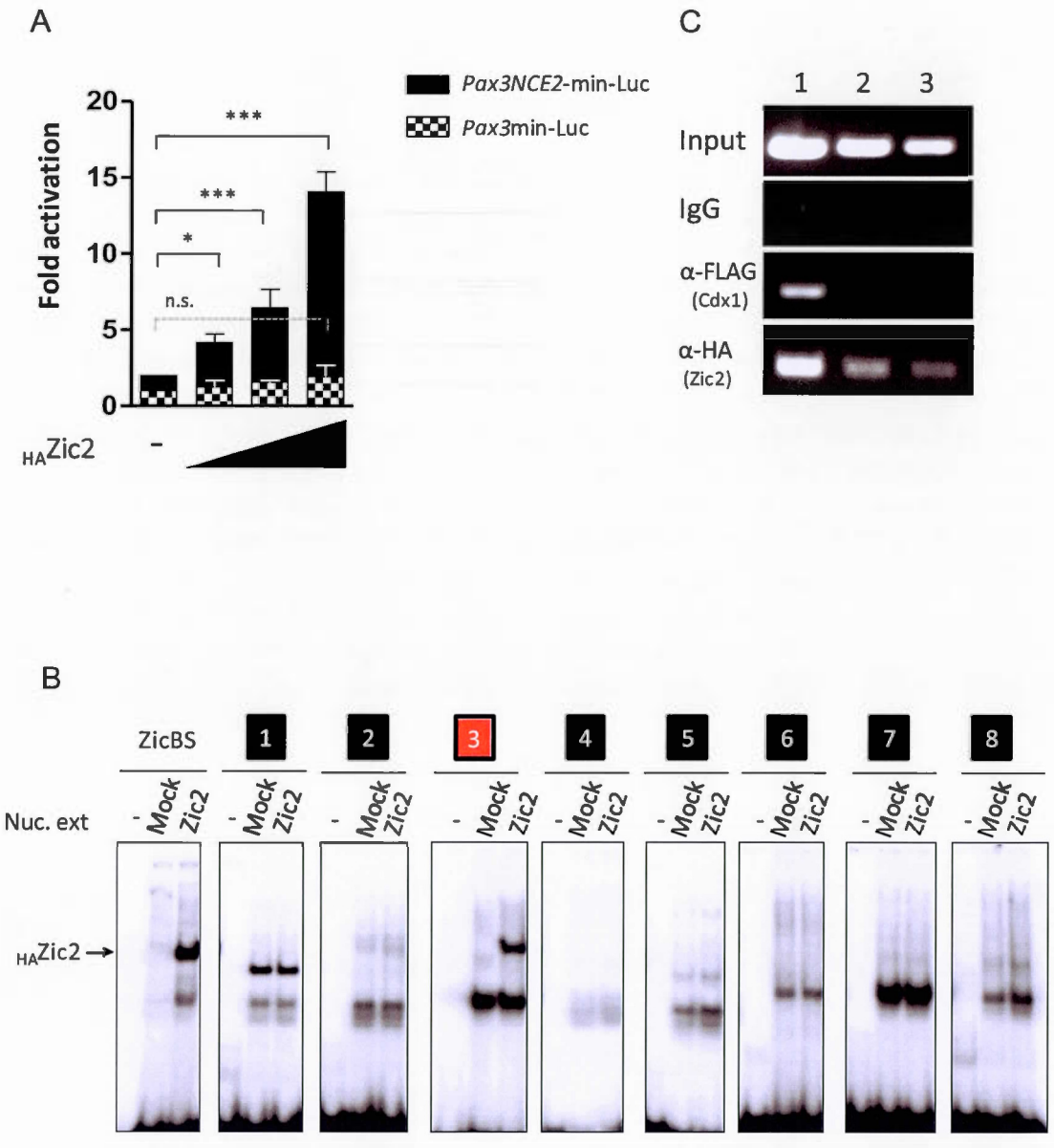


Figure 2.3 Zic2 activates and directly binds the *Pax3* neural crest enhancer NCE2.

(A) Co-transfection assays in N2a cells using luciferase reporter constructs driven by the *Pax3*NCE2 region plus the *Pax3* 150bp minimal promoter (*Pax3NCE2*-min-Luc) or only driven by the *Pax3* 150bp minimal promoter (*Pax3*min-Luc) as well as a $_{HA}$Zic2 expression vector. The luciferase quantification results are expressed as fold activation compared to each reporter vector alone. n=5 independent experiments performed in triplicate; one-way ANOVA: $P<0.0001$; Tukey's post test: n.s. $P>0.05$; * $P<0.05$; ***: $P<0.001$. Error bars

indicate s.e.m. Note that increasing Zic2 expression significantly increases the activation of the *Pax3*NCE2 reporter but not the *Pax3*min reporter. *(B)* Identification of Zic binding sites in the *Pax3*NCE2 by electrophoretic mobility shift assay (EMSA). Eight overlapping double-stranded oligonucleotide probes (represented as numbers 1 to 8 in boxes at the top of each panel) were used to scan the whole NCE2 for Zic binding sites (ZicBS). A probe bearing a consensus Gli/ZicBS (Mizugishi et al., 2001) was used as a positive control for Zic binding. All *in vitro* binding reactions were performed in parallel under identical conditions and each probe was either not incubated (-), or incubated with nuclear extracts from Mock- (transfected without DNA) or $_{HA}$Zic2-transfected Cos7 cells. Note that, similarly to the shifted band observed with the positive control (ZicBS), a specific shifted band is observed for radiolabelled probe #3 (red box) when incubated with $_{HA}$Zic2 nuclear extracts. *(C)* Chromatin immunoprecipitation assays in N2a cells showing the occupancy of *Pax3*NCE2 by both $_{FLAG}$Cdx1 and $_{HA}$Zic2. Anti-FLAG-, anti-HA- or normal rabbit IgG-immunoprecipitated chromatin extracts from $_{FLAG}$Cdx1- and $_{HA}$Zic2- transfected N2a cells were purified and amplified by PCR using primers flanking the *Pax3*NCE2 region. Amplification products, resolved on a 2% agarose gel, were of the expected size (246bp) and sequence-confirmed. Lanes 1, 2 and 3 represent PCR amplifications of serial dilutions of DNA. Note that *Pax3*NCE2 is amplified from chromatin samples immunoprecipitated with anti-FLAG (Cdx1) and anti-HA (Zic2), but not with the non-specific IgG.

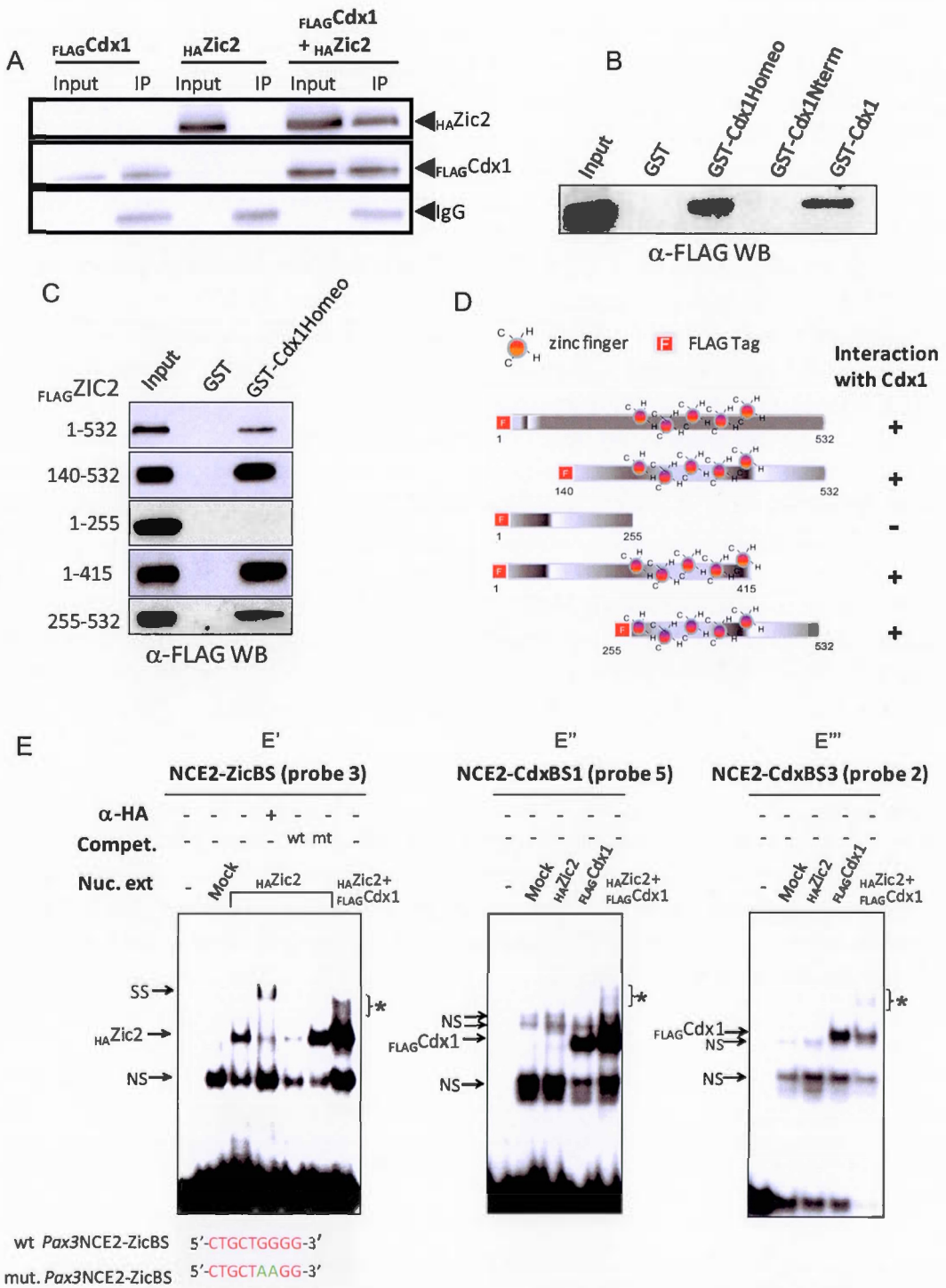


Figure 2.4 Zic2 and Cdx1 directly interact via their respective DNA binding domain.

(A) Co-immunoprecipitation assay using whole cell lysates from Cos7 cells transfected with $_{FLAG}$Cdx1 and/or $_{HA}$Zic2 expression vectors. Total lysates were immunoprecipitated with anti-FLAG antibody and analyzed by immunoblotting using an anti-HA antibody (top panel). Expression of Cdx1 was assessed by reprobing the blot with an anti-FLAG antibody (middle panel). Inputs represent 5% of the total lysate used for immunoprecipitation. *(B-C)* Mapping of Cdx1 and Zic2 interacting domains via GST pull-down assays. Inputs represent 15% of the total *in vitro*-translated proteins used for pull-downs. *(B)* Western blot using an anti-FLAG antibody showing that human $_{FLAG}$ZIC2 specifically interacts with both full-length Cdx1 (GST-Cdx1) and Cdx1 homeodomain (GST-Cdx1Homeo) but not with the Cdx1 N-terminal region (GST-Cdx1Nterm). *(C)* The Zinc finger domain of ZIC2 is essential for the interaction with the Cdx1 homeodomain. Different $_{FLAG}$ZIC2 deletion constructs were used to identify ZIC2 regions that interact with the Cdx1 homeodomain. Note that deletion of the five Cys2His2-type zinc fingers (construct 1-255) specifically abolishes interaction with the Cdx1 homeodomain. *(D)* Schematic representation of ZIC2 and its deletion constructs (adapted from (Pourebrahim et al., 2011) and (Houtmeyers et al., 2013)) and their interaction with Cdx1. *(E)* Characterization of the Zic binding site (ZicBS) contained in NCE2 and evaluation of the interaction of Zic2 and Cdx1 in the presence of their target DNA by EMSA. Results with a probe containing the new ZicBS (NCE2 scanning probe 3) are shown in E' while results with probes containing the previously described CdxBS1 (NCE2 scanning probe 5) or CdxBS3 (NCE2 scanning probe 2) are shown in E'' and E''', respectively. All *in vitro* binding reactions were performed in parallel under identical conditions. Binding of $_{HA}$Zic2 and $_{FLAG}$Cdx1 as well as anti-HA supershift (SS) and non-specific bindings (NS) are indicated by arrows. Brackets and asterisks denote the presence of higher molecular weight bands when $_{HA}$Zic2- and $_{FLAG}$Cdx1-overexpressing nuclear extracts are used in combination. Note that specificity of Zic2 binding to the new ZicBS was evaluated by pre-incubation of nuclear extracts from $_{HA}$Zic2- or mock-transfected Cos7 cells with a 100-fold molar excess of non-radiolabeled wild type (wt) or mutated (mt) probes. Sequences of the ZicBS as well as its mutated version used to assess Zic2 binding specificity by EMSA are at the bottom of E'; point mutations are denoted in green.

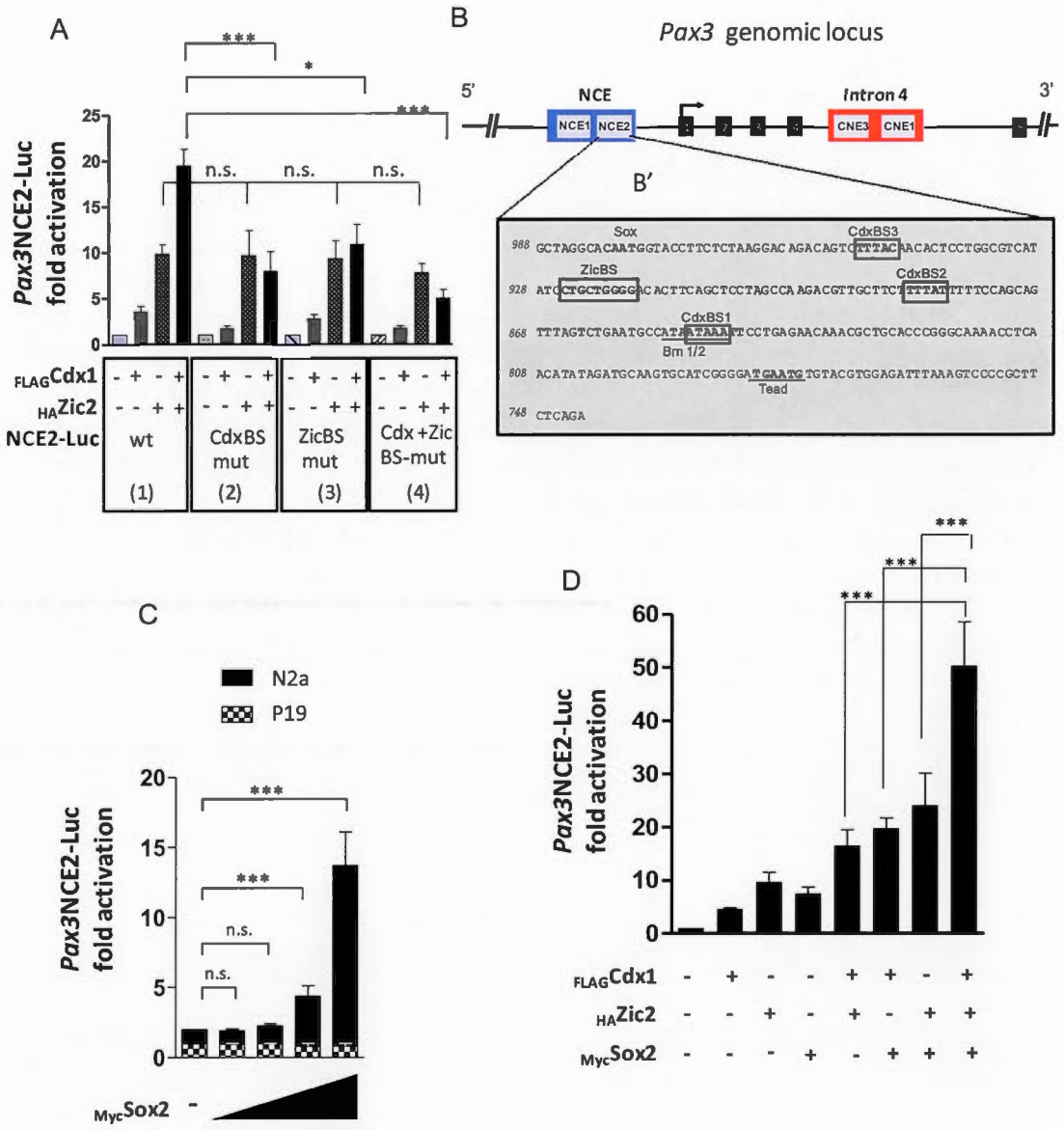


Figure 2.5 Cdx1 and Zic2 functionally interact and synergize with the SoxB1 family member Sox2 in the transactivation of *Pax3*NCE2.

(A) Cdx1 and Zic2 cooperatively transactivate *Pax3*NCE2 in N2a cells and Cdx binding sites (CdxBS) are essential for this functional interaction. A wild type (wt) *Pax3*NCE2-luciferase reporter (1) or mutant versions bearing mutations in the three CdxBS (2), the ZicBS (3) or a combination of these mutations (4) were evaluated for Cdx1 and Zic2 transactivation in N2a cells. Note that $_{FLAG}$Cdx1 and $_{HA}$Zic2 co-transfection results in a transactivation of the wt *Pax3*NCE2-luc reporter stronger than simple addition of each single transfections, and that mutation of CdxBS and/or ZicBS abrogates this cooperative effect. In contrast, the transactivating effect of Zic2 is not affected by mutation of CdxBS and/or ZicBS. *(B)*

Schematic representation of the *Pax3* genomic locus showing the relative position of *Pax3* regulatory regions. The 5' regulatory region (represented as a blue box) is named NCE and is subdivided in two CRMs called NCE1 and NCE2 (Milewski et al., 2004). The intron-4 regulatory region (red box) contains the CNE3 and CNE1 modules (Moore et al., 2013). Black boxes represent exons. *(B')* Magnification view of *Pax3*NCE2 DNA sequence showing the location of the ZicBS as well as the three functional CdxBS (in boxes) and in relation to a putative SoxBS (MatInspector analyses) as well as the previously described Brn1/2 (Pruitt et al., 2004) and Tead binding sites (Milewski et al., 2004) (underlined). *(C)* Sox2 strongly activates *Pax3*NCE2 in N2a cells, but cannot activate this CRM in undifferentiated P19 cells. N2a and P19 cells were equally transfected with the *Pax3*NCE2-luc reporter alone or with increasing amounts of the Sox2 expression vector. *(D)* Cdx1, Zic2 and Sox2 act synergistically in the transactivation of *Pax3*NCE2 in N2a cells. Cells were transiently transfected with the *Pax3*NCE2-luc reporter alone or co-transfected with $_{FLAG}$Cdx1, $_{HA}$Zic2 or $_{Myc}$Sox2 expression vector alone or in combination. In A, C and D, results are expressed as fold activation compared to the relevant reporter vector alone. n=6-10 independent experiments performed in triplicate; one-way ANOVA: $P<0.0001$; Tukey's post test: n.s. $P>0.05$; * $P<0.05$; ***: $P<0.001$. Error bars indicate s.e.m.

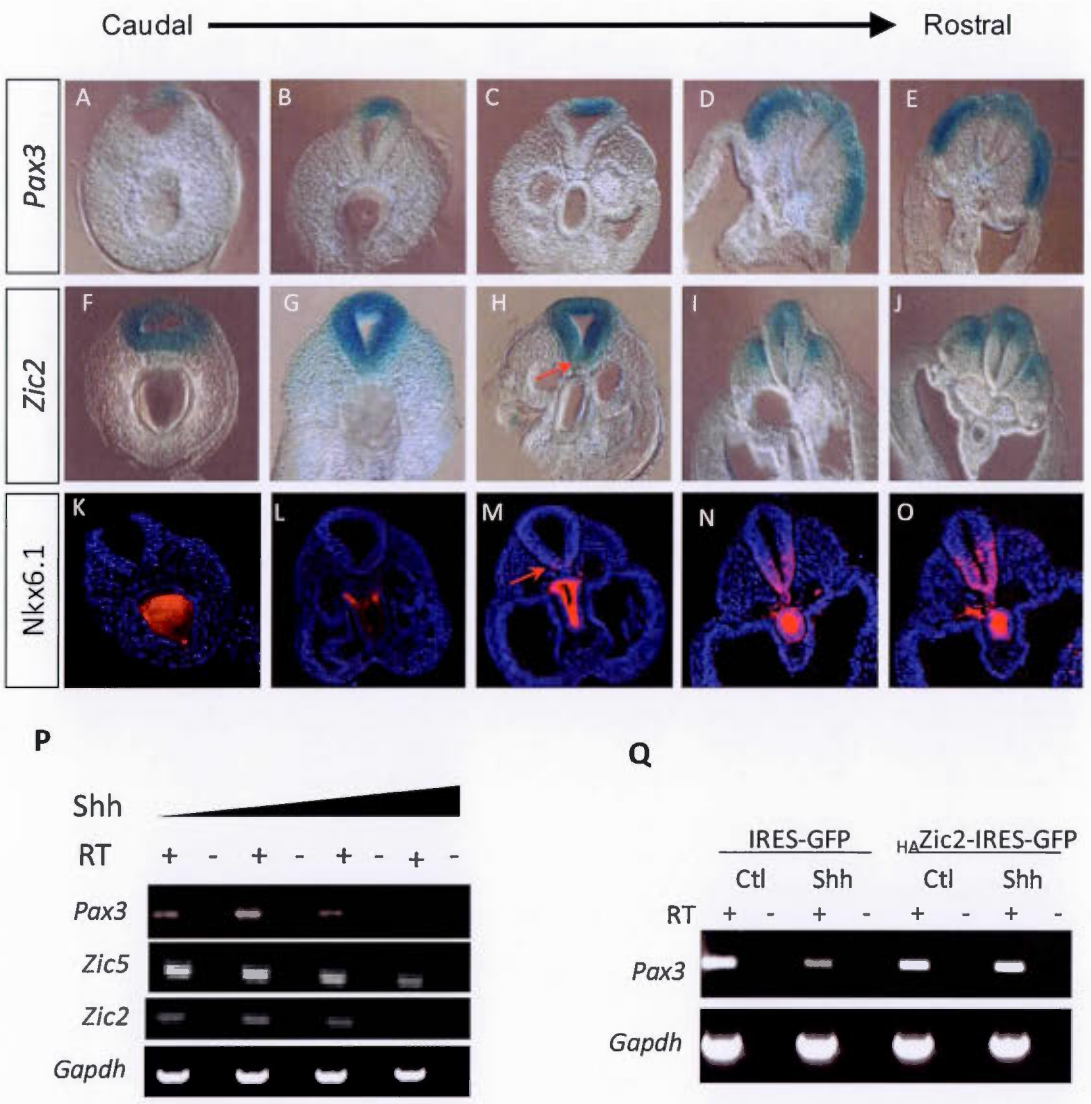


Figure 2.6 *Zic2* and *Zic5* are potential intermediates in the Shh-induced repression of *Pax3* expression.

(A-O) Comparison of *Pax3* and *Zic2* expression domains (assessed by whole-mount *in situ* hybridisation), with Nkx6.1 distribution (assessed by whole-mount immunofluorescence) in serial vibratome sections of the posterior end of 24-somite (e9.5) stage-matched embryos (caudal to rostral direction). Note that restricted *Pax3* expression in the open PNP (A) and recently closed NT (B) is established before the onset of Nkx6.1 expression in the ventral NT (red arrow in M). Also note that the onset of *Zic2* repression in the ventral NT (red arrow in H) correlates with the presence of Nkx6.1 in this region (red arrow in M). Going forward in the posterior to anterior direction, *Zic2* and Nkx6.1 expression domains are complementary

along the dorsal-ventral axis [i.e.: ~90% (H) vs ~10% (M); ~60% (I) vs ~40% (N); ~40% (J) vs ~60% (O)]. In contrast, *Pax3* and Nkx6.1 expression domains only become complementary at the most rostral levels [~40% (E) vs ~60% (O)]. *(P-Q)* Analysis of *Pax3, Zic2* and *Zic5* gene expression in Shh-treated N2a cells via semi-quantitative RT-PCR. Results shown are representative of at least two independent experiments. *(P)* The *Zic2-Zic5* gene pair is, like *Pax3,* expressed in N2a cells and negatively regulated by Shh in a dose-dependent manner. Prior to RT-PCR analysis, cells were cultured in absence or presence of increasing doses of Shh over a fixed 24-hour period of time. *(Q)* Zic2 is an intermediary in the mechanism of *Pax3* repression by Shh. Prior to RT-PCR analysis, cells were transfected with $_{HA}$Zic2 or empty IRES-GFP expression vector and then cultured for 24h in Ctl or Shh-enriched medium. Note that overexpression of Zic2 prevents the Shh-induced repression of *Pax3* in FACS-recovered N2a cells.

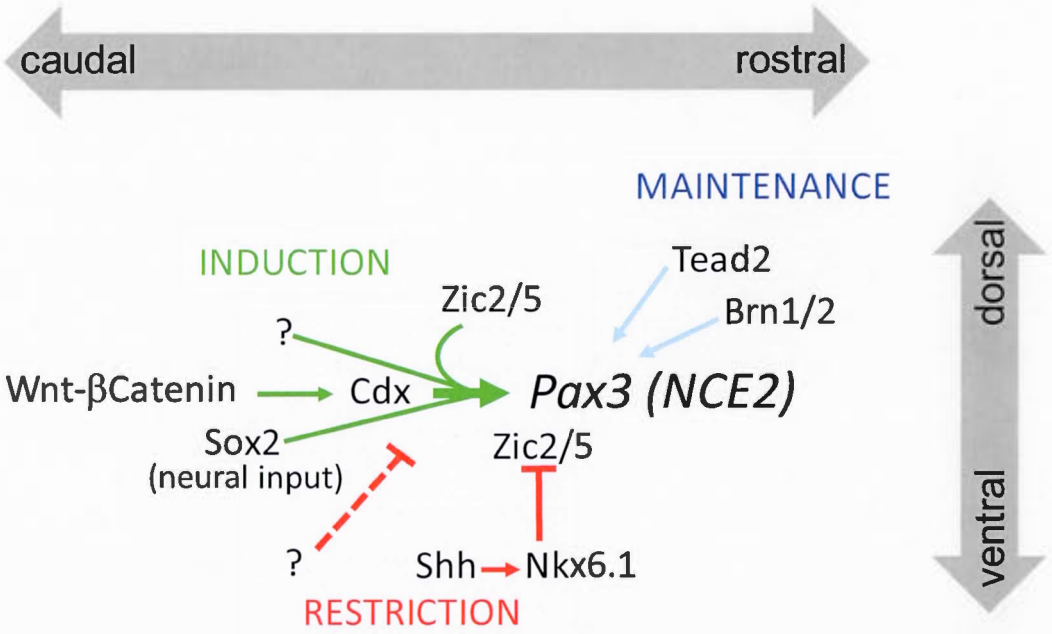


Figure 2.7 Current model for the Cdx-dependent control of *Pax3* expression in the caudal neuroectoderm via the NCE2 CRM.

Neural-specific induction of *Pax3* expression by the previously described Wnt-Cdx circuit requires an interaction with Zic2/5 and Sox2. Supplemental unknown (?) positive and/or negative inputs are required to restrict the *Pax3* expression domain to the lateral borders of the PNP. Maintenance of restricted NCE2 activity in the dorsal NT is most likely mediated by the dorsally restricted Zic2/5 transcription factors, which act as Cdx cofactors and are the predicted immediate early targets of the Shh-Nkx6.1 repression pathway. As elongation proceeds and the Cdx input disappears, *Pax3* expression in the NT is also later maintained by Tead2 and Brn1/2 transcription factors.

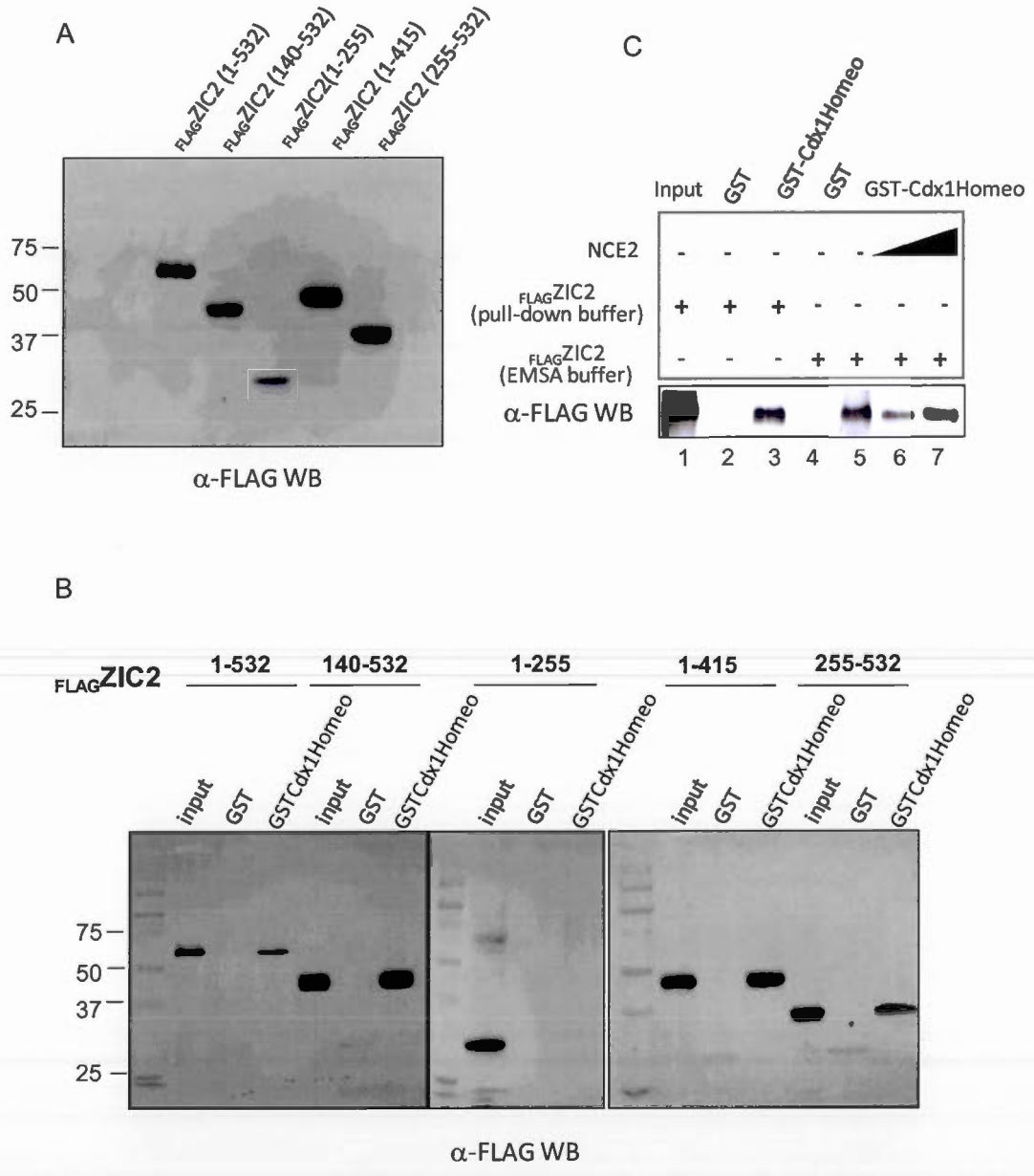


Figure 2.8 ZIC2 and Cdx1 can still physically interact in presence of the NCE2.

(A) Anti-FLAG western blot showing the expected size of *in vitro*-translated FLAG-tagged full-length ($_{FLAG}$ZIC2(1-532)) and truncated ZIC2 proteins ($_{FLAG}$ZIC2(140-532); $_{FLAG}$ZIC2(1-255); $_{FLAG}$ZIC2(1-415); $_{FLAG}$ZIC2(255-532)). *(B)* ZIC2 and Cdx1 bind to each other via their DNA binding domain. GST and GST-Cdx1 homeodomain (GST-Cdx1Homeo) fusion

proteins were bound to glutathione-agarose beads and then incubated with *in vitro*-translated full-length or truncated $_{FLAG}$ZIC2 proteins. Presence of $_{FLAG}$ZIC2 in the precipitated complexes was then analysed by Western blot using an anti-FLAG antibody. Inputs represent 15% of the total *in vitro*-translated proteins used for pull-downs. Note that specific bindings are observed with $_{FLAG}$ZIC2(1-532), $_{FLAG}$ZIC2(140-532), $_{FLAG}$ZIC2(1-415) and $_{FLAG}$ZIC2(255-532) proteins but not with the truncated version $_{FLAG}$ZIC2 1-255 lacking the zinc finger domain. *(C)* Anti-FLAG immunoblot analysis of the interaction between $_{FLAG}$ZIC2(1-532) and GST-Cdx1Homeo in presence of the target NCE2 DNA. Note that, in comparisons to standard pull-down conditions (lanes 1-3), using EMSA buffer conditions does not affect the interaction (lanes 4-5). Note also that variations in signal intensity detected in lanes 6-7 are independent of the amount of NCE2 sequences.

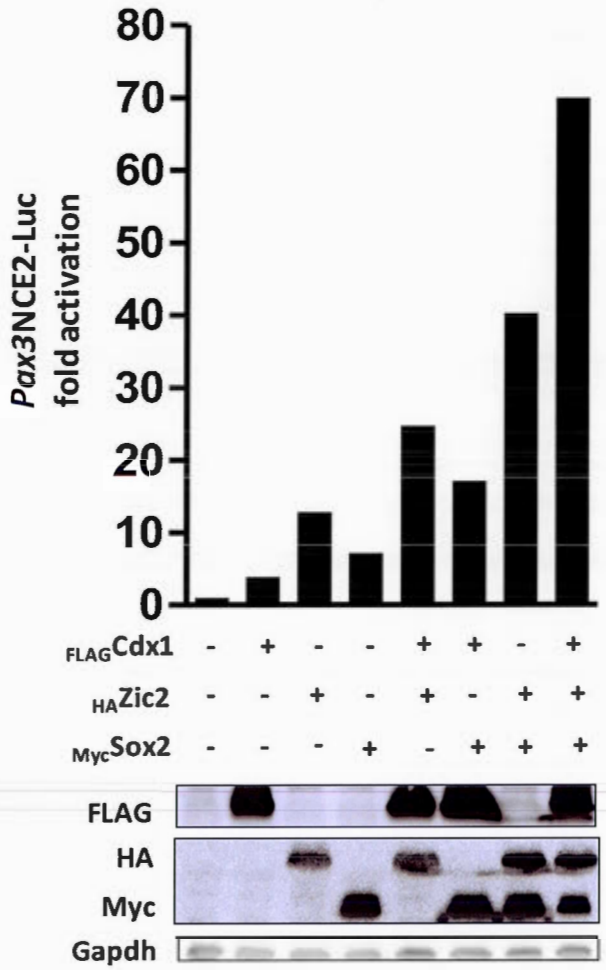


Figure 2.9 Western blot validation of the expression of Cdx1, Zic2 and Sox2 in N2a lysates used for luciferase reporter assays.

Expression vectors for $_{FLAG}$Cdx1, $_{HA}$Zic2 as well as $_{Myc}$Sox2 were transfected alone or in different combinations together with a luciferase reporter construct driven by the *Pax3NCE2* and the *Pax3* minimal promoter (*Pax3NCE2*-min-Luc). Production of exogenous proteins was evaluated by western blot using anti-FLAG, anti-HA as well as anti-Myc antibodies. Shown are the luciferase quantification results (expressed as fold activation compared to the reporter vector alone) as well as western blot validations for one representative experiment of the results displayed in Fig.5D.

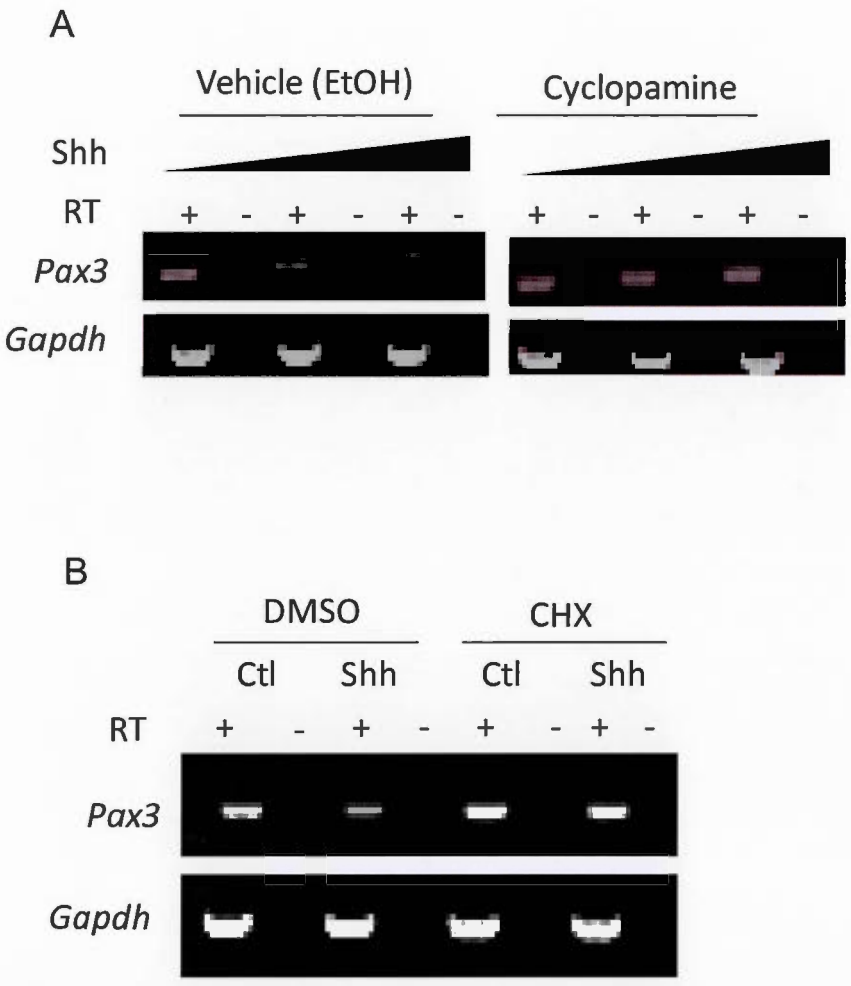


Figure 2.10 The N2a cell line is a good model for studying the Shh-induced repression of *Pax3* expression.

(A,B) RT-PCR analysis demonstrating that *Pax3* expression is indirectly repressed by Shh signaling in N2a cells. Prior to RT-PCR, cells were cultured for 24h in absence or presence of Shh and/or absence or presence of cyclopamine or cycloheximide (CHX). Results are representative of at least n=3 independent experiments. *(A) Pax3* expression is repressed by Shh in a dose-dependent manner (left panel) and this repression is abrogated in presence of cyclopamine, an inhibitor of the Shh receptor Smo (right panel). *(B) De novo* protein synthesis is required for the Shh-mediated repression of *Pax3* expression, as shown here in presence of the translation inhibitor CHX.

CHAPTER III

In vivo evidence for a novel and direct role for murine Caudal-related homeobox (Cdx) proteins in the trunk neural crest-gene regulatory network

(To be submitted to *Development*)

Oraly Sanchez-Ferras, Guillaume Bernas, Omar Farnos, Aboubacrine M. Touré and Nicolas Pilon.

Molecular Genetics of Development Laboratory, Department of Biological Sciences and BioMed Research Center, Faculty of Sciences, University of Quebec at Montreal (UQAM)

To whom correspondence should be addressed: Nicolas Pilon, Email: pilon.nicolas@uqam.ca

Author contribution:

Oraly Sanchez-Ferras: conception, design, acquisition, analysis and interpretation of data, drafting and revision of the article; (experiments: generation and characterization of the R26R-$_{FLAG}$EnRCdx1 mouse line; intercrossing of R26R$^{EnRCdx1/EnRCdx1}$ mice with *T*-Cre$^{Tg/+}$ mice and analysis of skeletal preparations; intercrossing of R26R$^{EnRCdx1/EnRCdx1}$ mice with *P3pro*-Cre$^{Tg/+}$ in WT, *Pax3*$^{+/Sp}$ and *Cdx1*KO backgrounds and characterization of NC and kidney phenotypes in embryos and adult mice (analysis of genetic interaction between R26(P3Cre)$^{EnRCdx1/+}$ with *Pax3*$^{Sp/+}$ and *Cdx1*KO mice, analysis of the expression of core NC regulatory genes in the mutants by *in situ* hybridization and identification of *Msx1* and *FoxD3* as novel Cdx targets; immunofluorescence analysis of c-Kit expression in R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO e11.5 embryos); generation of DNA constructs and design of co-transfection analysis in N2a cells; chromatin immunoprecipitation analysis in embryos; other preliminary data mentioned in the general discussion of the thesis. *Guillaume Bernas*: monitoring of mutant offspring, genotyping of mice and acquisition of data: analysis of the early postnatal lethality of R26(P3Cre)$^{EnRCdx1/+}$ mice, replication of *in situ* hybridization and immunofluorescence analysis, co-transfection analysis in N2a cells. *Omar Farnos*: monitoring of mutant offspring, genotyping of mice and acquisition of data: analysis of the early postnatal lethality of R26(P3Cre)$^{EnRCdx1/+}$ mice, replication

of *in situ* hybridization and immunofluorescence analysis, co-transfection analysis in N2a cells. *Aboubacrine M. Touré*: acquisition of data (whole mount acetylcholinesterase staining of colon preparations). *Nicolas Pilon:* conception, design, acquisition of data, supervision of the work and revision of the manuscript; (experiments: analysis of skeletal preparations).

3.1 Summary

Numerous studies in chordates and arthropods currently indicate that Cdx proteins have an ancestral major role in the organization of global posterior development (post-head tissues). One such study in the urochordate *Halocynthia roretzi* has notably reported that *Cdx* loss-of-function impairs axial elongation, neural tube (NT) closure and pigment cell development (Katsuyama et al., 1999). Intriguingly, in contrast to their key role in axial elongation and NT closure, a role for Cdx proteins in neural crest (NC)-derived melanocyte/pigment cell development has not been reported in any other chordate species. In this regard, although our recent work strongly supported the idea that murine *Cdx* genes (*Cdx1*, *Cdx2* and *Cdx4*) occupy a strategic position between NC inductive signals (Wnt/βCatenin) and a key NC regulator gene (*Pax3*), a global *in vivo* analysis of the Cdx role in trunk NC development was lacking. To address this, we generated a new conditional pan-Cdx functional knockdown mouse model that expresses a dominant negative fusion protein (EnRCdx1) in a Cre/loxP-dependent manner and thus circumvents Cdx functional overlap as well as the early embryonic lethality of *Cdx* mutants. Via targeted expression in the NC, we here provide *in vivo* evidence that murine Cdx proteins are required for melanocyte development and fulfill this role by directly controlling the expression of the key early regulators of the trunk NC-gene regulatory network *Pax3*, *Msx1* and *FoxD3*. Our work thus reveals a novel role for Cdx proteins atop of the trunk NC-gene regulatory network and melanocyte development in the mouse, which appears to be inherited from their ancestral orthologue.

3.2 Introduction

The neural crest (NC) is a vertebrate-specific transient population of multipotent cells that arise at a precise region of the embryo called the neural plate border (between the neural and non-neural ectoderm). Rolling-up of the neural plate during neurulation places the recently induced NC cells (NCC) in the dorsal part of the closing NT from where they eventually emigrate in order to colonize the embryo and differentiate into many derivatives (Bronner and LeDouarin, 2012). According to their location along the A-P axis, premigratory NCC are classified into cranial, vagal/cardiac, trunk and sacral populations (Bronner and LeDouarin, 2012). The trunk NC population has the unique potential to migrate either ventrally – to give rise to neural derivatives (such as neurons and glia of the dorsal root and sympathetic ganglia) – or dorso-laterally – to give rise to melanocytes of the skin.

Melanocytes are highly migratory melanin-producing cells that appear to represent the most ancient evolutionarily-conserved NC-like derivative. Such pigment cells have notably been found in urochordate embryos (Lamb et al., 2007; Nishida and Satoh, 1989; Yajima et al., 2003). Particularly, studies performed in the tunicate *Ciona intestinalis* have identified two melanocyte lineage precursos, the a7.6 cells and the a9.49 cells, which give rise to cephalic and trunk migratory cell populations exhibiting NC-like properties (Abitua et al., 2012; Jeffery, 2006; Jeffery et al., 2008; Jeffery et al., 2004).

NC development has been extensively studied in zebrafish, chick and frog, but much less in the mouse and human species (Betancur et al., 2010; Betters et al., 2010; Stuhlmiller and Garcia-Castro, 2012; Thomas et al., 2008). Evidence from these models and particularly from the study of the cranial NC population has led to propose a putative NC-gene regulatory network (GRN). This complex network notably involves the convergence of posteriorizing (canonical Wnt and FGF) and medio-lateral (BMP) signaling inputs at the neural plate border and the stepwise induction of first, a kernel of neural plate border specifier genes that encode

transcription factors such as Pax3/7, Msx1/2 and Zic1/Zic2, followed by a second module of NC specifier genes that encode transcription factors such as FoxD3 and the SoxE members Sox9/10 (Betancur et al., 2010). Comparative analysis of the cranial NC-GRN from basal chordates to vertebrate systems suggest that many regulatory circuits are conserved across chordates (Green and Bronner, 2013). However, it is known that the developmental behavior of trunk NCC largely differs from that of cranial NCC, and between species (Bronner and LeDouarin, 2012; Gammill et al., 2006; Pavan and Raible, 2012). Accordingly, both the manner by which the signaling inputs are integrated during trunk NC induction and the exact way by which the NC-GRN is wired at the transcriptional level are also expected to differ from the cranial region, and between species. How the trunk NC-GRN is wired in the mouse is largely unknown.

Abnormal NCC development results in numerous birth defects collectively known as neurocristopathies (Bolande, 1997), which generally involve pigmentary anomalies. Neurocristopathies are also frequently associated with NT defects (NTDs), which represent the second most prevalent group of human congenital malformations (Copp et al., 2003; Nye et al., 1999). Mutant mouse models have been particularly useful for finding genes with important role in NC and/or NT development in humans such as *PAX3*, which has been associated with *spina bifida* and pigmentation-related disorders (Waardenburg syndrome) in humans (Agopian et al., 2013; Auerbach, 1954; Baxter and Pavan, 2013). Moreover, overexpression of *PAX3* as well as that of other genes encoding important transcriptional regulators of melanocyte development such as *MITF*, *SOX10* and *FOXD3* have also been implicated in development of human melanoma, the most aggressive skin cancer (Levy et al., 2006; Shakhova et al., 2012; Weiss et al., 2014a). Understanding how the NC-GRN is wired in complex vertebrate models such as the mouse is thus necessary to identify new molecular effectors and to better understand the molecular mechanisms underlying these diseases.

The *Caudal-related homeobox (Cdx)* gene family encode homeodomain transcription factors with posterior-restricted expression patterns and key evolutionarily conserved roles in the general organization of post-head tissues (Copf et al., 2004; van Rooijen et al., 2012). Owing to the very high conservation of their helix-turn-helix homeodomain (Marom et al., 1997; Reece-Hoyes et al., 2002)(Marom et al., 1997; Reece-Hoyes et al., 2002)(Marom et al., 1997; Reece-Hoyes et al., 2002)(Marom et al., 1997; Reece-Hoyes et al., 2002)(Marom et al., 1997; Reece-Hoyes et al., 2002), murine Cdx proteins (Cdx1, Cdx2 and Cdx4) act redundantly in the regulation of their target genes (Savory et al., 2011a; Savory et al., 2009b; van den Akker et al., 2002; van Nes et al., 2006; van Rooijen et al., 2012). Between e7.5 and e12.5, robust expression of murine *Cdx* genes is initiated in epiblast cells undergoing gastrulation in the posterior regions of the embryo (i.e. primitive streak and axial stem cell zone of the tailbud). Their expression is then maintained in some derivatives of all three germ layers and most especially in the neuroectoderm, which represents the most anterior *Cdx* expression domain. At the onset of neurulation (e8.5), this is exemplified by a posterior (high) to anterior (low) gradient of transcript and protein distribution in the developing spinal cord, with an anterior limit of expression around the level of the hindbrain/spinal cord boundary set by Cdx1 (Beck et al., 1995; Gamer and Wright, 1993; Gaunt et al., 2003; Gaunt et al., 2005; Meyer and Gruss, 1993). Near this anterior limit, it is noteworthy that Cdx1 protein expression is limited to the dorsal NT and detected in NCC specifically migrating via the dorso-lateral pathway (Meyer and Gruss, 1993). Similarly, gene expression of both *Cdx2* and *Cdx4* has also been reported in the dorsal NT and NC lineages (Coutaud and Pilon, 2013a; Gaunt et al., 2005). The anterior limit of *Cdx* gene expression regresses caudally concomitantly with axial elongation and is maintained in the tailbud until E10.5 for *Cdx1* and *Cdx4* and E12.5 for *Cdx2*.

Cdx gene expression is regulated by posteriorizing signals from canonical Wnt, retinoic acid (RA) and FGF pathways, with all three *Cdx* being direct targets of

canonical Wnt signals in the mouse (Houle et al., 2000; Houle et al., 2003; Keenan et al., 2006; Lickert and Kemler, 2002; Pilon et al., 2006; Pilon et al., 2007; Prinos et al., 2001; Zhao et al., 2014). Cdx proteins are not only directly downstream of the canonical Wnt pathway, but can also modulate the Wnt input by directly interacting with the Lef/Tcf-ß-Catenin transcriptional complex (Beland et al., 2004; Grainger et al., 2013). The high expression level of *Cdx* genes in the caudal neuroectoderm during neurulation, together with their strategic position directly downstream of the Wnt/ß-Catenin input, suggest that Cdx proteins can mediate Wnt functions in early NC development.We previously tested this hypothesis and demonstrated that Cdx proteins indeed act downstream of canonical Wnt signaling to directly induce *Pax3* expression in premigratory NCC in cooperation with the neural plate border specifier Zic2 and the pan-neural factor Sox2 (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). However, the importance of this regulation for NC development *in vivo* was lacking.

Loss-of-function studies in the mouse have revealed that Cdx proteins control the A-P patterning of NC-derived spinal ganglia by regulating *Hox* gene expression (van den Akker et al., 2002). However, the understanding of the full scope of Cdx NC functions has been hampered by functional overlap as well as by the severe posterior truncation defects and early embryonic lethality of compound *Cdx*-null mutants (Savory et al., 2011a; van Rooijen et al., 2012). Here, in order to circumvent these limitations, we report the generation of a conditional (Cre-dependent) pan-Cdx loss-of function mouse model expressing a previously described Cdx dominant negative fusion protein (EnRCdx1; Sanchez-Ferras et al., JBC 2012) under the control of the *ROSA26* promoter. We directed expression of EnRCdx1 specifically in the dorsal NT and pre-migratory NCC using the *P3pro-Cre* line and show that Cdx proteins control melanocyte development in a cell-autonomous manner. We have characterized the molecular basis of this phenotype and found that Cdx proteins act early in the trunk NC-GRN by regulating expression of their known target *Pax3* as well as two other

important NC-regulator genes: *Msx1* and *FoxD3*. Our cis-regulatory analysis suggests that Cdx proteins regulate *Msx1* and *FoxD3* expression via binding to their proximal promoter, a common *modus operandi* of Cdx proteins. Our work thus places *Cdx* genes at the head of the trunk NC-GRN in the mouse.

3.3 Materials and Methods

3.3.1 Ethics Statement

Experiments involving mice were performed following Canadian Council of Animal Care (CCAC) guidelines for the care and manipulation of animals used in medical research. Protocols involving the manipulation of animals were approved by the institutional ethics committee of the University of Quebec at Montreal (*Comité institutionnel de protection des animaux* (CIPA)); Reference number 0513-C1-648-0514).

3.3.2 Plasmid constructs

Cdx1, Cdx2, Cdx4, and FLAG-tagged EnRCdx1 expression vectors co-expressing GFP (Cdx1-IRES-GFP; Cdx2-IRES-GFP; Cdx4-IRES-GFP or $_{FLAG}$EnRCdx1-IRES-GFP) were as previously described (Sanchez-Ferras et al., 2012) whereas the Cre-expressing vector pMC-Cre was kindly provided from D. Lohnes. For *in situ* hybridization probes, the *Pax3* cDNA (pBH3.2 clone) (Goulding et al., 1991) and HA-tagged Zic2 expression vector (pcDNA3-$_{HA}$Zic2) (Koyabu et al., 2001) were generous gifts from J. Epstein and J. Aruga, respectively. Plasmids containing the *FoxD3* cDNA (Hfh2) (Labosky and Kaestner, 1998)and *Msx1* cDNA (pSP72 Msx1) (Ishii et al., 2005) were kindly provided by KH Kaestner and R.E. Maxson Jr, respectively. *Sox9* (1.6kb) and *Sox10* (421bp) cDNA vectors were a kind gift from D.W. Silversides (Cory et al., 2007) whereas *Dct* (1.8kb) and *Mitf-M* (1.2kb) cDNA constructs were a generous gift from L. Sommer (Hari et al., 2012) (Bondurand et al., 2000).

For the generation of pROSA26-EnRCdx1 targeting construct, the $_{FLAG}$EnRCdx1-IRES-GFP sequence was excised from the EnRCdx1-IRES-GFP bicistronic construct with NheI and NotI and subcloned into the BigT vector (Srinivas et al., 2001), downstream of a *loxP*-flanked (PGKpNeo-tpA) stop cassette. The loxP-PGKpNeo-tpA-loxP-$_{FLAG}$EnRCdx1-IRES-GFP-bpA unit was then subcloned into the *Pac*I and *Asc*I restriction sites of pROSA26PA vector (Srinivas et al., 2001).

Msx1 and *FoxD3* 1.2 kb proximal promoter sequences were amplified by PCR from mouse genomic DNA, cloned in pGEM-T vector and validated by sequencing (primer sequences are available upon request). *Msx1*p1.2kb-Luciferase and *FoxD3*p1.2kb-Luciferase reporter constructs were then generated by subcloning *Msx1* or *FoxD3* proximal promoter sequences into pXP2 (Nordeen, 1988).

3.3.3 Generation of R26R-$_{FLAG}$EnRCdx1 "knock-in" mice

R1 embryonic stem (ES) cells were cultured as previously described (Pilon et al., 2007). These cells were electroporated with 30 μg of *Pvu*II-linearized pROSA26-EnRCdx1 targeting vector and selected with G418 (200 μg/ml) for 7 days. Surviving clones were isolated and analyzed for homologous recombination by genomic PCR and Southern blot. For PCR screening of G418-resistant ES clones, an external F1 forward primer (5'-AGGGCGGCTTGGTGCGTTTG-3') and an internal R1 reverse primer (5'-TGCGCCCTACAGATCCCTTA-3') were used to generate a 1.1 kb fragment. The genomic position of the primers is depicted in Figure 1B. Six PCR-positive clones were confirmed by Southern blotting using *Eco*RI and *Kpn*I double digested genomic DNA. Proper targeting was verified with a 550 bp 5'-external probe (11-kb wt allele and 4.0-kb targeted allele) as well as with a 800bp 3'-external probe (11 kb wt and 8.8 kb targeted allele), kindly provided by Christine Hartmann (Nyabi et al., 2009). Expression of the $_{FLAG}$EnRCdx1-IRES-GFP bicistronic transcript after removal of the PGKpNeo stop cassette was verified via anti-FLAG western blot analysis and GFP fluorescence following Cre transfection (pMC-Cre) of one targeted ES clone. This positive clone was used to generate germline chimeras by injection

into C57BL/6 blastocysts according to standard procedures (Nagy et al., 2003). F1 R26R$^{EnRCdx1/+}$ heterozygous mice were intercrossed to generate the R26R$^{EnRCdx1/EnRCdx1}$ homozygous line. Mice carrying the mutant allele were identified by PCR genotyping using primers flanking the ROSA26 *Xba*I integration site F2 (5'-CCGAAAGTCGCTCTGAGTTGTTATC-3') and R2 (5'-AGATGACTACCTATCCTCCCA-3') in combination with the R1 primer described above. The PCR product of R26R$^{EnRCdx1/+}$ heterozygous mice exhibited two bands, one of 345 bp in size corresponding to the wild type allele (F2/R2) and the other of smaller size (225 bp) corresponding to the mutant allele (F2/R1). R26R$^{EnRCdx1/EnRCdx1}$ homozygous mice were identified by the presence of only one band of 225 bp.

3.3.4 Mice

*Cdx1*KO mutant mice were generously provided by D. Lohnes with the permission of P. Gruss (Subramanian et al., 1995). *Pax3*$^{+/Sp}$ mice were obtained from Jackson laboratories (Bar Harbor, Maine). *T*-Cre mice were kindly provided by M. Lewandoski (Perantoni et al., 2005) whereas *P3pro*-Cre mice were kindly provided by S. Astrof with the permission of J. Epstein (Li et al., 2000). *P3pro*-Cre$^{Tg/+}$::*Pax3*$^{+/Sp}$ mice were generated by intercrossing *P3pro*-Cre$^{Tg/+}$ and *Pax3*$^{+/Sp}$ mice. The *P3pro*-Cre$^{Tg/+}$::*Cdx1*KO and R26R$^{EnRCdx1/EnRCdx1}$::*Cdx1*KO lines were generated by backcrossing the relevant allele into the *Cdx1*KO background. R26R$^{EnRCdx1/+}$::*T*-Cre$^{Tg/+}$ (renamed R26(TCre)$^{EnRCdx1/+}$) and R26R$^{EnRCdx1/+}$::*P3*pro-Cre$^{Tg/+}$ (renamed R26(P3Cre)$^{EnRCdx1/+}$) double heterozygous mice were generated by crossing R26R$^{EnRCdx1/EnRCdx1}$ mice with *T*-Cre$^{Tg/+}$ and *P3*pro-Cre$^{Tg/+}$ mice, respectively. R26R$^{EnRCdx1/+}$::*P3*pro-Cre$^{Tg/+}$::*Cdx1*KO (renamed R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO) mice or embryos were generated by crossing R26R$^{EnRCdx1/EnRCdx1}$::*Cdx1*KO and *P3*pro-Cre$^{Tg/+}$::*Cdx1*KO lines. R26R$^{EnRCdx1/+}$::*P3pro*-Cre$^{Tg/+}$::*Pax3*$^{+/Sp}$ (renamed R26(P3Cre)$^{EnRCdx1/+}$::*Pax3*$^{+/Sp}$) mice were generated by crossing R26R$^{EnRCdx1/EnRCdx1}$ mice with *P3pro*-Cre$^{Tg/+}$::*Pax3*$^{+/Sp}$ mice.

3.3.5 Offspring analysis

Embryos were generated by natural mating and collected between e9.5-e11.5, with noon of the day of vaginal plug detection considered as e0.5. Embryos to be compared were stage-matched according to established criteria (Kaufman, 1992) and processed in parallel. Whole mount *in situ* hybridization was performed as previously described (Hari et al., 2012; Ishii et al., 2005; Sanchez-Ferras et al., 2014) whereas transverse sections (100 μm) were prepared using a vibrating blade microtome Microm HM 650V (Thermo Scientific) as previously described (Coutaud and Pilon, 2013b). Whole mount skeletal preparations of newborns were performed as previously described (Allan et al., 2001). For acetylcholinesterase staining, the colon region from control and mutant P33 mice was dissected and stained in accordance with a previously described protocol (Enomoto et al., 1998). Images were acquired with a Leica DFC 495 camera mounted on a Leica M205 FA microscope (Leica Microsystems). Immunofluorescence on cryosections was performed as previously described (Sanchez-Ferras et al., 2014). Primary antibodies used were goat anti-Sox10 polyclonal antibody N-20 (1:100; Santa Cruz Biotechnology sc-17342) and goat anti-cKit polyclonal antibody (1:500; R&D systems AF1356). Alexa-647 and 594 secondary antibodies (1:500) were obtained from Jackson Immunoresearch. Images were taken with a Nikon A1 laser-scanning confocal microscope.

3.3.6 Chromatin immunoprecipitation (ChIP) assays

The trunk region from five freshly dissected e9.5 $R26R^{EnRCdx1/+}$::*P3pro*-$Cre^{Tg/+}$::*Cdx1KO* mouse embryos was isolated and cross-linked with 1% formaldehyde in PBS for 15 min at room temperature. ChIP assays were then performed using the M-Fast Chromatin immunoprecipitation kit (ZmTech Scientific) in accordance with manufacturer's instructions. Immunoprecipitation was performed as previously described (Sanchez-Ferras et al., 2014) using 1 ug of mouse anti-FLAG M2 (Sigma) antibodies. Rabbit IgG was used as a negative control of immunoprecipitation. PCR amplifications were performed using the Taq Hifi DNA

polymerase (Invitrogen) and consisted of 35 cycles of 30s at 94°C, 30s at 60°C and 90s at 68°C. The primers used for amplification of a 276 bp evolutionarily conserved region (ECR) from the *FoxD3* proximal promoter (position -1058 to -782 bp relative to the transcription start site) were: *FoxD3*-Forward (5'-CTCCGTTTCCAGCTCATTTCCGAC-3') and *FoxD3*-Reverse (5'-CTCAAGTCTTACCCTAGCTTTCCG-3') primers. Primers for amplification of a 813 bp region (position -1137 bp to -325 relative to the transcription start site) of the *Msx1* proximal promoter, containing several ECR, were: *Msx1*-Forward (5'-CTCCCCAACAGCCTGTTCGAAC-3') and *Msx1*-Reverse (5'-CCCACTTCTTCTGTTCCTCTCC-3'). PCR products were resolved on a 2 % agarose gel and confirmed by sequencing.

3.3.7 Transfection analysis

Neuro2a (N2a) cell line propagation, transfection and luciferase reporter assays were performed as previously described (Sanchez-Ferras et al., 2012). For analyzing the effect of Cdx expression on the activity of the *Msx1* and *FoxD3* promoters, N2a cells ($3x10^4$) were seeded in 24-well plates and transfected with 100 ng of respective luciferase reporter construct together with increasing amounts of Cdx1-IRES-GFP, Cdx2-IRES-GFP or Cdx4-IRES-GFP expression vectors (from 0 to 20 ng). An empty IRES-GFP expression vector was also included to complete the final amount of DNA per well to 125 ng. For Cdx and EnRCdx1 competition assays, N2a cells were similarly transfected with 100 ng of the respective luciferase reporter construct but with a fixed amount of Cdx1 (5ng), Cdx2 (5ng) and Cdx4 (10 ng) expressing vectors (alone or in combination) and increasing amounts of EnRCdx1-IRES-GFP expression vector (from 0 to 30 ng). An empty expression vector was also included to complete the final amount of DNA per well to 150 ng, when necessary. All transfections were performed at least three times in triplicate.

3.4 Results

3.4.1 Generation of a mouse model allowing conditional pan-Cdx functional knockdown

To study the Cdx role in NC development *in vivo*, we devised a strategy aimed at allowing a neuroectoderm-specific pan-Cdx functional knockdown. This approach is based on the Cre/loxP-dependent expression of a FLAG-tagged fusion protein ($_{FLAG}$EnRCdx1) consisting of the repressor domain of *Drosophila* Engrailed (EnR) fused to the Cdx1 homeodomain. Successful usage of such Cdx obligatory repressor *in vivo* has been previously reported for several species (Bel-Vialar et al., 2002; Isaacs et al., 1998; Katsuyama et al., 1999; Mita and Fujiwara, 2007). It is also important to note that the Cdx homeodomain is highly conserved between Cdx1, Cdx2 and Cdx4 proteins and is not only used for DNA target recognition but also for mediating protein-protein interactions (Beland et al., 2004; Lafontaine et al., 2012; Marom et al., 1997; Sanchez-Ferras et al., 2014). Therefore, via the Cdx1 homeodomain, EnRCdx1 is able to compete with endogenous Cdx proteins for Cdx binding sites (CdxBS) as well as for interacting partners. Importantly, both the efficacy and specificity of EnRCdx1 as a repressor of Cdx target genes regulated either by DNA binding-dependent (via CdxBS) or -independent means (via an interaction with a DNA binding co-factor such as Lef1; (Beland et al., 2004)) has been previously validated (Sanchez-Ferras et al., 2012).

As depicted in Figure 3.1A, a $_{FLAG}$EnRCdx1-IRES-GFP bicistronic cassette preceded by a floxed-PGKpNeo cassette was targeted to the *ROSA26* locus by homologous recombination in ES cells, thereby generating a novel *ROSA26* reporter allele that we have named R26R^{EnRCdx1}. In addition to having served for ES clone selection purposes, the floxed-PGKpNeo cassette here serves as a transcriptional stop that prevents expression from the ubiquitously active *ROSA26* promoter. Specific Cre-mediated removal of this interfering cassette and expression of the $_{FLAG}$EnRCdx1-IRES-GFP module was validated in one targeted ES cell clone by anti-FLAG western blot and visualization of GFP fluorescence (Figure 3.1B,C). Standard microinjection

of these targeted ES cells in blastocysts yielded three germ-line transmitting chimeras. PCR-genotyped $R26R^{EnRCdx1/+}$ heterozygous mice were recovered at the expected mendelian ratio and intercrossed to obtain $R26R^{EnRCdx1/EnRCdx1}$ animals. These mice were also recovered at the expected frequency, were healthy and fertile, and were intercrossed to generate the $R26R^{EnRCdx1/EnRCdx1}$ line.

3.4.2 *T* promoter-directed expression of EnRCdx1 recapitulates the vertebral patterning defects of *Cdx* compounds mutants

To validate the efficacy and specificity of the conditional EnRCdx1 mouse model, we first evaluated whether it would allow affecting the well characterized Cdx role in A-P vertebral patterning (Chawengsaksophak et al., 1997; Lohnes, 2003; Pilon et al., 2007; Savory et al., 2009a; Savory et al., 2009b; Subramanian et al., 1995; van den Akker et al., 2002; van Nes et al., 2006). To this end, $R26R^{EnRCdx1/EnRCdx1}$ mice were crossed with *T*-$Cre^{Tg/+}$ mice (Perantoni et al., 2005), which express the Cre recombinase in the posterior nascent mesoderm under the control of *brachyury* (*T*) promoter (Figure 3.10A).

$R26R^{EnRCdx1/+}$::*T*-$Cre^{Tg/+}$ mutants ($R26(TCre)^{EnRCdx1/+}$) were obtained at the expected mendelian ratio and did not exhibit any overt phenotype. However, skeletal preparations of $R26(TCre)^{EnRCdx1/+}$ neonates showed the presence of vertebral defects extending from the cervical down to the lumbar regions of the axial skeleton (Figure 3.2). A high number of vertebral malformations was observed at the cervical level, with most of them being fully penetrant (Figure 3.2A and D). Cervical vertebras (C) were frequently anteriorly transformed, fused and/or exhibited malformed neural arches, suggesting defects in vertebral patterning and somitogenesis. All mutants presented the atlas or C1 vertebra fused with basioccipital bones concomitant with a C2 to C1 transformation. C3 to C2 transformations were also observed but with a lower penetrance (67%). The anterior arch of the atlas (AAA) was malformed and frequently fused to basioccipital bones (67%) and the tuberculum anterior (TA) on C6 was absent in 100% of mutants. Other, less frequent defects were observed in the

compact and malformed cervical region of R26(TCre)$^{EnRCdx1/+}$ neonates and are summarized in Figure 3.2D.

Skeletal malformations were also observed at the thoracic and lumbar regions, but with incomplete penetrance (Figure 3.2B and C). In some mutants, the thoracic and lumbar regions were shorter, as indicated by the presence of a total of 12 instead of 13 thoracic vertebras in 17% of the mutants (Figure 3.2B'), and a total of 5 instead of 6 lumbar vertebras in 33% of the mutants (Figure 3.2C). Fusion of ribs was also observed in the thoracic region (Figure 2B''), with a penetrance similar to that observed in *Cdx1*$^{-/-}$/*Cdx2*$^{+/-}$ offspring (33%) (van den Akker et al., 2002). In summary, these data show that single allelic expression of EnRCdx1 in the mesoderm generate axial skeleton malformations similar to those observed in *Cdx1*$^{-/-}$*Cdx2*$^{+/-}$ compound mutants. This clearly validates the R26R^{EnRCdx1} mouse model as a potent genetic tool for conditional *Cdx* loss-of-function studies.

3.4.3 *Pax3*promoter-directed expression of EnRCdx1 results in pigmentation defects and hydronephrosis

To knockdown Cdx function in the neuroectoderm we used the *P3pro*-Cre line, which expresses the Cre recombinase in the dorsal NT and NC under the control of the *Pax3* 1.6 kb proximal promoter (Li et al., 2000). Following mating of R26R$^{EnRCdx1/EnRCdx1}$ with *P3pro*-Cre mice, we first validated the occurrence of the Cre recombination event in resulting R26(P3Cre)$^{EnRCdx1/+}$ e9.5 embryos using GFP fluorescence as a surrogate marker for EnRCdx1 expresssion (Figure 3.10). The fluorescence pattern recapitulated the *lacZ* expression pattern reported for e9.5 transgenic embryos carrying *P3pro*-lacZ constructs (Natoli et al., 1997). GFP fluorescence was broadly distributed in the tailbud region and appeared progressively more neuroectoderm-specific anteriorly (Figure 3.10 B, C and D). Transverse section at the level of the forelimb buds showed high levels of GFP expression in cells emigrating from the dorsal NT (Figure 3.10E) and co-expression with Sox10 confirmed that these GFP-positive cells are *bona fide* NCC (Figure 3.10F,G). Weak

GFP expression was also observed throughout the entire dorso-ventral axis of the NT although it was found to be stronger in the dorsal region (Figure 3.10E). This pattern of Cre activity has already been reported in other studies using the *P3Pro*-Cre line (Liu et al., 2006a).

Initial analysis of 7 breeding couples indicated that R26(P3Cre)$^{EnRCdx1/+}$ offspring are recovered at lower than expected mendelian ratio at weaning (14% instead of 50%, Table 3.3). Careful monitoring of the offspring indicated that 72% of mutant mice die before P10. A high percentage of these mice (65%) did not survive more than 2 days, and lacked milk in the stomach indicating abnormal feeding. Importantly, 100% of recovered R26(P3Cre)$^{EnRCdx1/+}$ mutant mice exhibited pigmentary anomalies. Lack of pigmentation was observed in the hind paws and distal tail in all mutants, but never in anterior regions of the body (Figure 3.3D, Table 3.1). In some cases, a very small white spot was also observed on the belly (Figure 3.4A). To determine whether these pigmentation defects were specifically caused by reduction in Cdx function, we expressed EnRCdx1 in a *Cdx1*KO background. Decrease in *Cdx* dosage significantly increased the extent of white spotting, thereby phenocopying the *Pax3* Splotch heterozygous (*Pax3*$^{Sp/+}$) phenotype in posterior regions (Figure 3.4). Such posterior restriction is in total agreement with the restricted expression of *Cdx* genes in the posterior embryo and is indicative of a new cell-autonomous role for Cdx proteins in trunk NC development.

Other observations revealed that recovered R26(P3Cre)$^{EnRCdx1/+}$ mutants also exhibit severe growth delay (Figure 3.3A). Intriguingly, 26% of these mutants were found dead within 3 weeks and dissection revealed the presence of a hydronephrotic kidney in most of the cases (data not shown). Systematic analysis of a random group of R26(P3Cre)$^{EnRCdx1/+}$ mutants revealed that this hydronephrosis phenotype is in fact present in only about half of cases (53%), in an age-independent manner, and frequently found to unilaterally affect the left kidney (Figure3.3B and Table 3.1). Hydronephrosis is thus unlikely to represent the major cause of death of the born

mutants. Presence of a mega-bladder was also observed in some cases but with a lower penetrance (13%) relative to the hydronephrosis phenotype (data not shown and Table 3.1). Moreover, a kinked tail was occasionally noted and analysis of an additional random group of R26(P3Cre)$^{EnRCdx1/+}$ mutants revealed that this phenotype is present in 44% of cases (Figure 3C and Table 3.1). As previously reported for some *Cdx* gain- and loss-of function mouse models (Gaunt and Paul, 2011; van de Ven et al., 2011; van Rooijen et al., 2012), it is also noteworthy that R26(P3Cre)$^{EnRCdx1/+}$ mutant mice did not reproduce well and frequently exhibited purulent abscesses in the genital and anal regions. Unfortunately, this prevented the generation of R26(P3Cre)$^{EnRCdx1/EnRCdx1}$ animals (Table 3.3).

3.4.4 R26(P3Cre)$^{EnRCdx1/+}$ genetically interacts with *Pax3*$^{Sp/+}$ in the development of NC-derived melanocytes and enteric nervous system

Our previous work underscored a key role for Cdx proteins in the induction of *Pax3* expression in trunk NCC via direct binding to a short evolutionary-conserved enhancer (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). To expand on this *in vivo* – and based on the presence of *Pax3*$^{Sp/+}$-like pigmentation defects in R26(P3Cre)$^{EnRCdx1/+}$ mutants (Figure 3.4) – we asked whether NC-directed expression of EnRCdx1 in a *Pax3*$^{Sp/+}$ background would yield significant reduction of *Pax3* expression to increase the pigmentation defects of *Pax3*$^{Sp/+}$ mutants. To this end, R26R$^{EnRCdx1/EnRCdx1}$ and *P3pro*-Cre$^{Tg/+}$::*Pax3*$^{Sp/+}$ mice were mated in order to generate *R26R*$^{EnRCdx1/+}$, R26R$^{EnRCdx1/+}$::*Pax3*$^{Sp/+}$, R26(P3Cre)$^{EnRCdx1/+}$ and R26(P3Cre)$^{EnRCdx1/+}$::*Pax3*$^{Sp/+}$ allelic combinations (Figure 3.5). Analysis of phenotype and mendelian ratios of these animals revealed that the lethality (Table 3.3) as well as the postnatal growth delay and pigmentary anomalies observed in R26(P3Cre)$^{EnRCdx1/+}$ mice were all significantly accentuated in presence of a *Pax3*^{Sp} allele (Figure 3.5 G,H). As *Pax3* function is also essential for the formation of the enteric nervous system, we then evaluated whether the R26(P3Cre)^{EnRCdx1}::*Pax3*^{Sp} genetic interaction would be in addition reflected by defective rostro-caudal colonization of the bowel by NC-derived enteric neural progenitors as observed in

Pax3$^{Sp/Sp}$ mutants (Lang et al., 2000). Since the colon is the last segment to be colonized by NC-derived enteric neural progenitors, the integrity of the enteric nervous system was specifically analyzed in this region via staining of acetylcholinesterase activity in muscle strips (i.e. only containing the myenteric plexus) (Figure 3.5 I-L). This analysis revealed that, while the colon of *Pax3*$^{Sp/+}$ mutants is fully colonized, a discrete aganglionic zone is observed in the most posterior regions of the colon in R26(P3Cre)$^{EnRCdx1/+}$ mutants and the extent of this aganglionic zone is exacerbated in R26(P3Cre)^{EnRCdx1}::*Pax3*^{Sp} (Figure 3.5 K-M). In summary, these results show that *EnRCdx1* and *Pax3* genetically interact in the development of two NC-derived lineages: melanocytes and enteric neural progenitors.

3.4.5 EnRCdx1-mediated Cdx loss-of-function affects both the number and the location of melanoblasts

In order to begin understanding the mechanism by which Cdx regulate melanocyte development, we analyzed the expression of the melanoblast marker cKit receptor tyrosine kinase, (Steel et al., 1992) in stage-matched control (R26R$^{EnRCdx1/+}$::*Cdx1*KO) and R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO e11.5 embryos (Figure 3.6). Immunofluorescence analysis for cKit revealed a surprising increase in the number of melanoblasts migrating along the dorso-lateral pathway at the level of the hindlimb buds in the mutants (Figure 3.6). Intriguingly, some of these cKit+ melanoblasts were also found to be ectopically located in the dorsal root ganglia (DRG) (Figure 3.6A). The presence of melanoblasts in ectopic location is indicative of a defect in the ability of these cells to normally respond to their environmental cues and thus makes defective guidance a likely cause of the pigmentation defects observed in EnRCdx1 mutants.

3.4.6 Expression of the core NC-regulatory genes *Pax3*, *Msx1*, *FoxD3* and *Sox9* is deregulated upon EnRCdx1-mediated Cdx loss-of-function

Post-head NCC are constantly induced at the lateral borders of the posterior neural plate, concomitantly with axial elongation and generation of new neural cells from

the posterior growth zone. These newly induced NCC are first marked by the expression of a set of genes – including *Pax3*, *Msx1* and *Zic2/5* – that are known as the neural plate border specifiers (Catron et al., 1996; Houzelstein et al., 1997; Merzdorf, 2007; Sanchez-Ferras et al., 2014). As the embryo elongates and neurulation goes on, pre-migratory NCC then become localized to the dorsal part of the NT and now also express genes such as *FoxD3*, *Sox9* and *Sox10,* which are known as the NC specifiers. In an effort to identify Cdx NC targets, we evaluated the expression of these early markers of the trunk NC-GRN via whole mount ISH in control ($R26R^{EnRCdx1/+}$::*Cdx1*KO) and mutant $R26(P3Cre)^{EnRCdx1/+}$::*Cdx1*KO e9.5 embryos (Figure 3.7).

The analysis of neural plate border specifiers first revealed that expression of *Pax3* in pre-migratory NCC is reduced in mutant embryos, from the forelimb buds down to the tailbud (Figure 3.7 A-B,A'-B'). Of note, *Pax3* somitic expression was found to remain unchanged which is in total agreement with our previous work showing that Cdx proteins regulate *Pax3* expression via a trunk NC-specific enhancer (*Pax3*NCE2) (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012) (asterisk in Figure 3.7A', B'). Interestingly, expression of the neural plate border specifier *Msx1* in premigratory NCC was also found to be significantly reduced from the forelimb buds down to the tailbud region in $R26(P3Cre)^{EnRCdx1/+}$::*Cdx1*KO mutants (Figure 3.7 C-D;C'-D'). This EnRCdx1-mediated repressive effect was again found to be neural specific, as *Msx1* expression in the limb buds was not affected (asterisk in Figure 7 C'-D'). On the other hand, the mRNA levels of the Cdx co-factor Zic2 (Sanchez-Ferras et al., 2014) were apparently not affected in mutants compared to controls (Figure 3.7E-F, E'-F').

Similar to *Pax3* and *Msx1*, the transcript levels of the NC specifier *FoxD3* were also found to be significantly reduced in pre-migratory and early migratory NCC of $R26(P3Cre)^{EnRCdx1/+}$::*Cdx1*KO embryos (Figure 3.8 A-B; A'-B'). In marked contrast, *Sox9* expression was found to be increased in premigratory trunk NCC of mutant

embryos (E-F;E'-F'). Again, the unaffected *Sox9* expression in somitic tissues is indicative of a NC-specific effect (red asterisk in inset Figure 8 E',F'). Finally, *Sox10* expression levels were not significantly affected in premigratory NCC of mutant embryos (Figure 3.8 C-D; C'-D').

Taken together, the reduction of *Msx1* and *FoxD3* transcript levels in premigratory and early migratory trunk NCC upon Cdx loss-of-function suggests that these genes represent novel positive Cdx targets. The observed upregulation of *Sox9* in mutant NCC may be an indirect consequence due to the down-regulation of a *Sox9* repressor or compensation, since EnRCdx1 acts as an obligate repressor.

3.4.7 Cdx proteins transactivate the proximal promoter of *Msx1* and *FoxD3*

As several studies have shown, Cdx members often regulate expression of their direct target genes by binding to their proximal (<2kb) promoter region (Charite et al., 1998; Sanchez-Ferras et al., 2012; Savory et al., 2011a; Subramanian et al., 1995). Consequently, to understand the molecular mechanism by which Cdx proteins activate *Msx1* and *FoxD3* expression, we first looked for Cdx binding sites in the proximal promoter region of these genes. In both cases, a considerable enrichment for consensus-like Cdx binding sites was found in the 1200 bp region upstream of the transcriptional start site (Figure 3.9A,B,C). As verified using the ECR browser, these promoter regions are highly conserved between mouse and human (Figure 3.9A,B), and previous studies have also shown that they are sufficient for activating reporter gene expression (Kuzuoka et al., 1994; Takahashi et al., 1997).

We then generated luciferase reporter constructs driven by these proximal promoter sequences (*Msx1*p1.2kb-Luc; *FoxD3*p1.2kb-Luc) and assessed the role of the Cdx proteins in their regulation via co-transfection analyses in the NC-derived Neuro2a cell line (Figure 3.11 and 3.9D). This analysis revealed that every Cdx member can transactivate in a dose-dependent manner both the *Msx1* and the *FoxD3* proximal promoters (Figure 3.11). *Cdx2* overexpression gave the highest transcriptional response, followed by *Cdx1* and *Cdx4* (Figure 3.11), consistent with the known

differential transactivity of Cdx proteins (Grainger et al., 2013; Sanchez-Ferras et al., 2012). Interestingly, co-transfection of all three Cdx proteins synergistically activated the *Msx1* and *FoxD3* proximal promoters (Figure 3.9D) and this effect was decreased in a dose-dependent manner by $_{FLAG}$EnRCdx1 (Figure 3.9D).

To confirm occupancy at the endogenous loci, we performed ChIP-PCR assays using trunk tissues from R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO e9.5 embryos. Importantly, this showed that $_{FLAG}$EnRCdx1 binds *in vivo* to conserved and CdxBS-enriched sequences within the *Msx1* and *FoxD3* proximal promoters (Figure 3.9E). These results are in accordance with unpublished anti-Cdx2 ChIP-seq data showing peaks located at -647 bp and -823 bp from the *Msx1* and *FoxD3* transcription start site, respectively (D. Lohnes, personal communication). Taken together, these results demonstrate that Cdx proteins directly regulate *Msx1* and *FoxD3* expression via occupancy and transactivation of their proximal promoter.

3.5 Discussion

In spite of the high *Cdx* expression in the caudal neuroectoderm across chordates, the Cdx role in the NC has remained obscure because of functional overlap and early embryonic lethality of *Cdx* mutants. We previously demonstrated that Cdx proteins are at the heart of a regulatory circuit that control the induction and dorsal restriction of *Pax3* expression in the mouse caudal neuroectoderm (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). To validate the functional importance of this regulation for NC development *in vivo*, and unravel new *Cdx* roles in this "fourth germ layer", we needed to circumvent functional overlap, early embryonic lethality and severe posterior truncation phenotypes normally associated with single and compound *Cdx* mutants (Savory et al., 2011a; van Rooijen et al., 2012). We thus generated a new Cre/LoxP-dependent pan-Cdx loss-of-function mouse model that conditionally expresses a constitutive repressor form of Cdx (EnRCdx1) (Sanchez-Ferras et al., 2012) and used the P3proCre line to express EnRCdx1 in the dorsal NT and NC with an anterior limit in the caudal hindbrain from e8.5 onwards. Single allelic expression

of EnRCdx1 in these neural tissues resulted in posterior-specific pigmentary anomalies, which recapitulate the *Pax3* Splotch heterozygous phenotype. Consistent with our previous mechanistic studies we observed a reduction in *Pax3* transcript levels in the caudal neuroectoderm of R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO e9.5 embryos. Our study thus confirms that a Wnt-Cdx-Pax3 pathway is at work *in vivo* and supports a novel role for Cdx proteins in the control of melanocyte development in the mouse.

3.5.1 *Cdx* genes at the head of the trunk NC-GRN

The NC developmental program begins with the specification of the neural plate border by the integration of posteriorizing (canonical Wnt, FGF and RA) and medio-lateral (BMP) instructive cues and the resulting induction of *Msx1/2*, *Pax3/7* and *Zic* genes (Betancur et al., 2010). Previous work from our group and others led us to hypothesize that, at least in the mouse, *Cdx* genes are at the head of the trunk NC-GRN, downstream of these inductive cues. Indeed, *Cdx* genes are known to be regulated by these signaling pathways (Lengerke et al., 2008; Lohnes, 2003; Pilon et al., 2006; Pilon et al., 2007; Prinos et al., 2001) and are expressed in the caudal neuroectoderm before the neural plate border specifiers (Catron et al., 1996; Goulding et al., 1991; Meyer and Gruss, 1993; Nagai et al., 1997). In addition, Cdx proteins are also known to mediate the regulation of canonical Wnt and BMP target genes by directly interacting with the respective nuclear effectors of these signaling pathways, the Lef-Tcf/ß-Catenin complex (Beland et al., 2004) and pSmad1/5/8 (Mari et al., 2014). On the other hand, we have shown that Cdx proteins integrate canonical Wnt signaling to notably induce *Pax3* expression in the caudal neuroectoderm via direct binding to a neural crest enhancer (Sanchez-Ferras et al., 2012) and in direct collaboration with the neural plate border specifier Zic2 (Sanchez-Ferras et al., 2014). The current work now adds *Msx1* and *FoxD3* as novel direct Cdx target genes and thus reveals that Cdx-dependent regulatory circuits can be reiteratively used to control many steps of early trunk NC development: in

specification of the posterior neural plate border via regulation of *Pax3* and *Msx1* and in specification of the NC via regulation of *FoxD3*. Altogether, these results thus clearly support a key *Cdx* role at the head of the trunk NC-GRN in the mouse.

Since *Cdx* gene expression is restricted to the caudal embryo, our work further confirms that the NC-GRN is dynamic and differentially wired along the A-P axis (Bronner and LeDouarin, 2012). Together with the well-known and major Cdx role in the control of A-P patterning, this also suggests that multifunctional Cdx-dependent regulatory circuits can allow coordinating NC specification with NC positional identity along the A-P axis. Indeed, downstream of posteriorizing signals, Cdx can integrate both inductive information via regulation of key neural plate border and NC specifier genes (this work) and positional information via regulation of multiple *Hox* genes (Bel-Vialar et al., 2002; Nordstrom et al., 2006) . Interestingly, such role does not appear to be restricted to *Cdx* genes since *Gbx2* and *Hoxb5* – two other A-P patterning genes – have also been reported to play important early roles in the NC-GRN (Kam et al., 2014a; Li et al., 2009).

3.5.2 Cdx proteins and the early control of trunk NC development

The spatiotemporal distribution of Cdx proteins in the mouse embryo is consistent with a role for these transcription factors in the control of the early steps of trunk NC ontogenesis, including induction, survival, proliferation and delamination (Coutaud and Pilon, 2013a; Meyer and Gruss, 1993). Accordingly, our previous and present work revealed that Cdx proteins directly upregulate the expression of the key early regulators of NC development *Msx1*, *Pax3* and *FoxD3*. Of note, we also detected an increase in *Sox9* transcripts levels in pre-migratory NCC of R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO embryos. Since EnRCdx1 acts as a repressor, this upregulation must be due to the down-regulation of a *Sox9* repressor for which FoxD3 is a good candidate. Indeed, FoxD3 is a known transcriptional repressor (Sutton et al., 1996) and it has been shown that *FoxD3* loss-of-function in the NC

lineage (NC-*FoxD3* cKO) similarly leads to an upregulation of *Sox9* expression in mouse embryos (Mundell and Labosky, 2011).

The reduced expression of *Pax3* observed in R26(P3Cre)$^{EnRCdx1/+}$ embryos, as well as the presence of *Pax3*$^{Sp/+}$-like phenotypes in postnatal animals (pigmentation and enteric nervous system defects) on either a WT or a *Cdx1*KO genetic background and the genetic interaction revealed on a *Pax3*$^{Sp/+}$ background are indicative of a problem with the development of NC progenitors at least downstream of *Pax3*. Interestingly, *Pax3* is upstream of *FoxD3* in the NC-GRN and both genes play similar critical early roles in the control of maintenance of the trunk NC progenitor pool. Indeed, *Pax3*$^{Sp/Sp}$ homozygous mutants or NC-*FoxD3* cKO embryos exhibit a severe reduction in the number of NC progenitors, which consequently leads to absence of NC derivatives including enteric ganglia and melanoblasts in the case of *Pax3*$^{Sp/Sp}$ mutants (Epstein et al., 2000; Hornyak et al., 2001; Lang et al., 2000; Olaopa et al., 2011; Teng et al., 2008). Accordingly, a likely possibility for explaining the observed NC defects in the EnRCdx1 model was that the number of neural and melanogenic NC-progenitors required for colonizing the skin and the intestine is reduced. Indeed, the observed aganglionosis in the distal colon of R26(P3Cre)$^{EnRCdx1/+}$ mice and the increase in severity in a *Pax3* Splotch heterozygous background sustain this hypothesis. Howeversuch a reduction in the number of NC progenitors in R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO embryos may be below the detectable levels since we failed to detect a difference in the number of NCC compared to control, either at e9.5 or e11.5 according to immunofluorescence analysis for the NC marker Sox10 (data not shown). This lack of effect can also be explained by the fact that single allelic expression of EnRCdx1 is not enough to lead to a complete loss of *Pax3* and *FoxD3* expression and therefore can only allow recapitulating the mild *Pax3*$^{Sp/+}$ phenotypes. Alternatively, *Sox9* upregulation may have also compensated for the loss of NC progenitors since it has been observed that ectopic *Sox9* expression in the chick NT leads to increased NC specification (Cheung and Briscoe, 2003). Moreover,

reduced *Msx1* expression is also expected to contribute to the NC defects of R26(P3Cre)$^{EnRCdx1/+}$. Indeed, *Msx1* loss-of-function studies in *Xenopus* have been shown to result in a lack of pigmentation concomitant with reduced expression of *Pax3* and *FoxD3* (Monsoro-Burq et al., 2005; Simoes-Costa et al., 2012). However, such a *Msx1* role in the mouse has not been reported so far, probably because of early perinatal lethality of *Msx1/2* single and compound mutants and functional overlap with *Msx3* (Ishii et al., 2005; Satokata and Maas, 1994).

3.5.3 Cdx proteins and the molecular control of melanocyte development via Pax3, FoxD3 and Sox9

Pax3, FoxD3 and Sox9 are all known essential regulators of melanocyte development, although they play opposing roles in this lineage. On one hand, analysis of the *Pax3*$^{Sp/Sp}$ melanoblast population has revealed that an early role of Pax3 is to control the number of melanocyte precursors (Hornyak et al., 2001). Pax3 promotes melanoblast specification by activating the expression of the master melanocyte regulator Mitf isoform M and also maintains these cells in an undifferentiated state by repressing activation of the *Dct* promoter by Mitf (Kubic et al., 2008; Lang et al., 2005). In contrast, FoxD3 prevents melanoblast specification by inhibiting the binding of Pax3 to the *Mitf* promoter (Thomas and Erickson, 2009) and thereby controls the timing of melanocyte specification (Nitzan et al., 2013). FoxD3 also controls melanoblast pathfinding as indicated by the premature and transient entrance of these cells into the dorso-lateral pathway as well as their ectopic presence in the dorsal NT, dorsal root and sensory ganglia of NC-*FoxD3* cKO embryos (Nitzan et al., 2013). On the other hand, perturbation analyses in chick embryos have shown that Sox9 can either promote or repress melanogenesis in a stage-dependent manner. *Sox9* overexpression in the dorsal NT (i.e. at an early stage) induces melanocyte differentiation including premature entrance into the dorsolateral pathway (Cheung and Briscoe, 2003), whereas ectopic *Sox9* expression in already specified melanoblasts (i.e. at a later stage) represses melanogenesis and dorso-lateral pathfinding (Nitzan et al., 2013). The latter is also consistent with the

hypopigmentaion phenotype observed in Ods heterozygous mouse mutants as a result of ectopic and persistent *Sox9* expression in melanoblasts from e11.5 onwards (Qin et al., 2004).

Taking into account the aforementioned observations, our results showing an increased number of R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO melanoblasts migrating into the dorsolateral pathway as well as their ectopic presence in the dorsal root ganglia are therefore much more reflective of the observed *FoxD3* downregulation and *Sox9* upregulation rather than the *Pax3* downregulation. On the other hand, although abnormal melanoblast pathfinding might contribute to the observed pigmentation defects of R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO animals, this phenotype obviously cannot be explained by an increase number of melanoblasts migrating along the normal dorso-lateral pathway. These apparent conflicting observations clearly highlight the need for an in-depth analysis of the fate of the mutant melanoblasts. Future analysis will also be needed to characterize the particular contribution of *Pax3*, *FoxD3*, and *Sox9* to the pigmentary anomalies observed in EnRCdx1 mutants.

3.5.4 Direct regulation of *Msx1* and *FoxD3* expression by Cdx proteins

The molecular mechanisms that control *FoxD3* and *Msx1* expression in the trunk NC are not well understood and most especially in the mouse. Studies in *Xenopus* place *Msx1* upstream of both *Pax3* and *FoxD3* in the NC-GRN (Monsoro-Burq et al., 2005; Simoes-Costa et al., 2012) whereas studies in chick and mouse embryos place *FoxD3* downstream of the redundant paralogs *Pax3/7* (Dottori et al., 2001; Simoes-Costa et al., 2012). Given the cross-regulatory interactions between these core NC-GRN genes, we cannot exclude the possibility that, for example, the decreased *FoxD3* expression in R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO e9.5 embryos might be an indirect consequence of a decreased expression of its regulators Pax3 and Msx1. However, our ChIP and promoter analyses indicate that Cdx proteins occupy and transactivate the CdxBS-enriched proximal promoter of both *Msx1* and *FoxD3*. This work thus extends the list of Cdx target genes that are regulated via their proximal promoter

region (Charite et al., 1998; Mari et al., 2014; Sanchez-Ferras et al., 2012; Savory et al., 2011a; Subramanian et al., 1995). Since Cdx proteins are restricted to the caudal embryo, our results also mean that expression of *Foxd3* and *Msx1* is regulated by region-specific mechanisms along the A-P axis.

It is noteworthy that like Pax3, both *Msx1* and *FoxD3* are downstream of canonical Wnt signaling (Bang et al., 1999; Sato et al., 2005; Taneyhill and Bronner-Fraser, 2005; Zhao et al., 2014). Our previous and present data showing that Cdx regulate the proximal promoters of *Pax3*, *Msx1* and *Foxd3*(Sanchez-Ferras et al., 2012), together with the central position of Cdx directly downstream of the canonical Wnt input (Pilon et al., 2006; Pilon et al., 2007; Sanchez-Ferras et al., 2012; Zhao et al., 2014) and known role as integrator of this posteriorizing signal (Lohnes, 2003), suggest that Cdx concomitantly relay the Wnt inductive input to several promoters in order to activate the NC-developmental program. Transgenic reporter analysis in the mouse has identified a small CRM 78 bp long called proximal enhancer (PE) which is located 2.6 kb upstream of the *Msx1* transcription start site (TSS) and recapitulates reporter expression in the roof plate and other regions of E10.5 transgenic embryos (MacKenzie et al., 1997). Whereas PE has been shown to bear functional TCF and Smad binding sites (Alvarez Martinez et al., 2002; Miller et al., 2007), it is unclear whether this CRM drives induction of *Msx1* in the posterior neural plate border. On the other hand, studies in *Xenopus* and chick have shown that induction of *FoxD3* requires stimulation with canonical Wnt signaling, but the molecular mechanism of this regulation are not well understood (Sato et al., 2005; Taneyhill and Bronner-Fraser, 2005). Our work constitutes the first to propose how canonical Wnt may indirectly control *FoxD3* induction in the trunk NC population via Cdx and activation of the proximal promoter. Cdx are known modifiers of chromatin configuration (Verzi et al., 2013) and we cannot rule out the possibility that via binding to the proximal promoter Cdx coordinate chromatin looping and enhancer-promoter interaction. In support of this hypothesis, it has been shown that PE activity in the

dorsal NT of transgenic embryos is strongly reduced in the absence of the *Msx1* proximal promoter region. In the case of *FoxD3*, cis-regulatory analyses in chick have identified a distal NC enhancer named NC2 located approximately 44 kb upstream of the TSS that recapitulates induction of FoxD3 in trunk NCC (Simoes-Costa et al., 2012). It is also probable that via binding to the *Foxd3* proximal promoter Cdx mediate long-range interaction with the distally located trunk NC enhancer NC2.The broad distribution of canonical Wnt and Cdx proteins in the posterior neural plate at the onset of NC induction, and then in the whole D-V axis of the recently closed NT, implies that the Wnt-Cdx pathway must cooperate with other factors to restrict *Msx1* and *FoxD3* expression to the NC domain. Consistent with our work, a recent study has shown that the A-P patterning protein Hoxb5 controls NCC survival via direct binding and regulation of the *FoxD3* proximal promoter (Kam et al., 2014a). As Cdx proteins have been shown to physically and functionally interact with several Hox homeoproteins including members of the Hox5 group (Lafontaine et al., 2012), it is thus possible that Cdx proteins cooperate with Hoxb5 in the transactivation of the *FoxD3* proximal promoter. However, like Cdx proteins, Hoxb5 is broadly expressed in the caudal neuroectoderm, meaning that other yet unknown factors restrict the Cdx input to the NC domain. In the case of *Msx1*, sequence analysis suggest that the 1.2 kb proximal promoter of *Msx1* contains enough cis-regulatory information to integrate positional cues from BMP, canonical Wnt, and FGF signaling (Alvarez Martinez et al., 2002; Bushdid et al., 2001; Kuzuoka et al., 1994). *Msx1* is notably a well known target of BMP signaling for which intermediate levels have been proposed to be involved in the restriction of *Msx1* expression to the neural plate border region (Liem et al., 1995; Tribulo et al., 2003). Moreover, this regulation of *Msx1* by BMP signaling appears to be mediated by the binding of pSmad1/5/8 proteins to a consensus binding site contained in the proximal promoter (Alvarez Martinez et al., 2002). Given the recent description of Cdx-pSmad1/5/8 physical interactions in intestinal tissues (Mari et al., 2014), it is thus tempting to

hypothesize that such a transcriptional complex also operates on the *Msx1* proximal promoter in order to restrict *Msx1* expression to the neural plate border region.

3.5.5 An ancestral role for Cdx proteins in pigment cell development

Our work provides the first *in vivo* evidence for a novel and direct *Hox*-independent role for Cdx proteins in pigment cell development in the mouse. Nevertheless, it is noteworthy that a lack of pigmentation has been previously reported following Cdx loss-of-function in the ascidian *Halocynthia roretzi* (Katsuyama et al., 1999). Intriguingly, Cdx-dependent NT defects have been also exclusively reported in ascidian and mouse embryos and not in any other chordate species (Katsuyama et al., 1999; Mita and Fujiwara, 2007; Savory et al., 2011a). Whether these Cdx ancestral functions in pigment cell and NT development have re-emerged in the mouse or have been evolutionarily-conserved but masked by functional overlap and/or early embryonic lethality in other chordate species is currently unknown. In support of the former, it is important to note that Cdx loss-of function in zebrafish negatively affects the number of Rohon-Beard mechanosensory neurons, which share a common origin with NCC (Epperlein et al., 2007; Skromne et al., 2007). Noteworthy also, the two NC-like body pigment lineage precursos identified in the ascidian specie *Ciona intestinalis* are regulated by canonical Wnt signaling, as well as express homologs of several vertebrate NC-regulatory genes such as *FoxD3* (Abitua et al., 2012; Jeffery et al., 2008). Taking into account our previous work, showing that vertebrate Cdx proteins are integrators of canonical Wnt signals during NC ontogenesis (Sanchez-Ferras et al., 2012), the present work showing that *FoxD3* is a Cdx target gene and the reports of pigmentation defects upon Cdx loss-of-function in one ascidian specie, these observations suggest that *Caudal-related* genes might be amongst the most ancestral components of the NC-GRN.

3.6 Conclusion

This study provides the first evidence of a direct role for Cdx proteins in trunk NC and melanocyte development in the mouse. Our analysis place *Cdx* genes at the top of

the trunk NC-GRN downstream of the inductive inputs and directly upstream of *Pax3*, *Msx1* and *FoxD3*.

3.7 Acknowledgements

The authors thank Denis Flipo (UQAM) for FACS analyses and assistance with confocal imaging as well as Qinzhang Zhu and Li Lian (Institute de recherches cliniques de Montréal) for microinjections. David Lohnes (University of Ottawa), Mark Lewandoski (National Cancer Institute, Maryland), Jonathan Epstein (Penn Cardiovascular Institute), Sophie Astrof (Thomas Jefferson University), David W. Silversides (Faculté de médecine vétérinaire, Université de Montréal), Sabine Tejpar (Katholieke Universiteit Leuven), Jun Aruga (RIKEN Brain Science Institute), Klaus Kaestner (University of Pennsylvania), Robert E. Maxson (University of Southern California Keck School of Medicine), Lucas Sommer (University of Zurich) and Christine Hartmann (Institute of Molecular Pathology, Vienna) are thanked for kindly agreeing to provide mice and/or reagents. David Lohnes is also thanked for having agreed to share unpublished information regarding Cdx2 ChIP-seq data. We thank Lieven Haenebalcke (Ghent University) and Jody Haigh (Monash University) for their advice regarding southern blot analysis of ES clones. Melanie Beland (Institute for research in immunology and cancer) is thanked for her advice during generation of R26R-$_{FLAG}$EnRCdx1 "knock-in" mice. This work was supported by a grant from the Canadian Institute for Health Research (CIHR grant number MOP-111130) to NP. OSF was supported by an Alexander-Graham-Bell scholarship from the Natural Science and Engineering Research Council (NSERC) of Canada whereas NP is a Fonds de la Recherche du Québec – Santé (FRQS) Junior2 scholar.

3.8 Figures

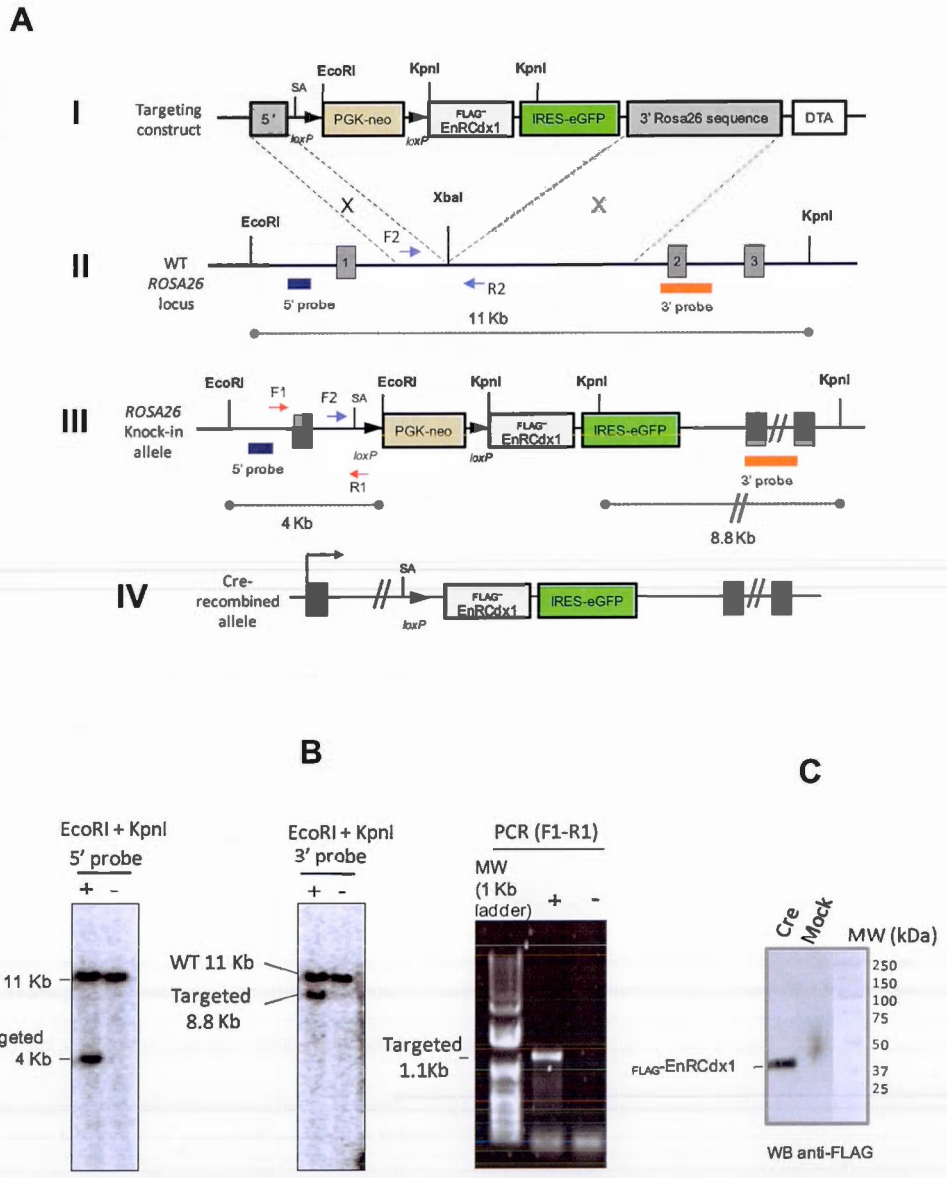
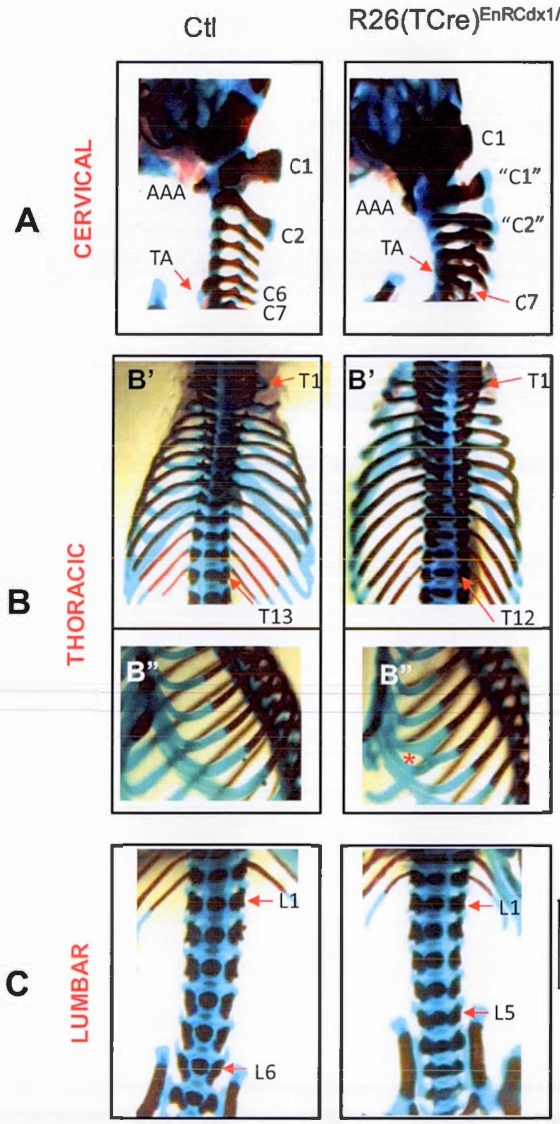


Figure 3.1 Targeting of dominant negative $_{FLAG}$EnRCdx1 coding sequences in the ROSA26 locus by homologous recombination in ES cells.

(A) Schematic representation of: (I) the targeting construct, (II) the wild type ROSA26 locus, (III) the targeted allele containing a bicistronic cassette coding for FLAG-tagged EnRCdx1 and GFP downstream of a floxed-PGKpNeo cassette, as well as (IV) the Cre-recombined allele where the floxed PGKp-Neo cassette is excised and expression of EnRCdx1-IRES-GFP is driven by the ROSA26 promoter. Blue and orange boxes represent the 5' and 3' external probes, respectively, used to screen for correctly targeted ES clones by Southern blot. Red arrows represent the F1 and R1 primers used for the screening of ES clones by PCR, whereas blue arrows depict the F2 and R2 primers used in combination with the R1 primer to genotype $R26R^{EnRCdx1}$ mice ROSA26 exons 1, 2 and 3 are denoted in gray boxes. The dashed lines denote the expected sizes of genomic DNA after double digestion with EcoRI and KpnI and hybridization with the 5' and 3' probes: wild type allele (11 kb using 5' or 3' probes), targeted allele (4kb, using the 5'probe) or (8.8 kb, using the 3'probe). *(B)* Southern blot and PCR confirmation of one correctly targeted ES clone used to generate $R26R^{EnRCdx1/+}$chimeras; analysis of a negative clone (-) is also included for matters of comparison. *(C)* Confirmation by Western blot anti-FLAG of the Cre-mediated expression of $_{FlAG}$-EnRCdx1 in the targeted ES clone. Prior to WB analysis, targeted ES cells were transfected with Cre-expressing pMC-Cre (Cre) or empty (mock) expression vectors.



D

CERVICAL REGION	
PHENOTYPE	**No. (%)**
Basioccipital fusion to AAA	4/6 (67%)
Vertebra 1	
Malformed NA	6/6 (100%)
Fusion to basioccipital	6/6 (100%)
Absence of AAA	1/6 (17%)
Vertebra 2	
C1 identity	6/6 (100%)
Fusion to C1	5/6 (83%)
Malformed NA	6/6 (100%)
Vertebra 3	
C2 identity	4/6 (67%)
Malformed NA	4/6 (67%)
Fusion to C4	3/6(50%)
Vertebra 4	
Fusion to C3	3/6(50%)
NA fusions	
C1-C2	5/6 (83%)
C3-C4	3/6 (50%)
Absence of TA (vertebra 6)	6/6 (100%)
Malformed vertebra 7	1/6 (17%)

THORACIC REGION	
PHENOTYPE	**No. (%)**
12 thoracic vertebra instead of 13	1/6 (17%)
Fusion of ribs	2/6 (33%)

LUMBAR REGION	
PHENOTYPE	**No. (%)**
5 lumbar vertebra instead of 6	2/6 (33%)

Figure 3.2 Conditional expression of $_{FLAG}$EnRCdx1 in the mesoderm recapitulates the A-P vertebra patterning defects of Cdx mutants.

Skeletal analysis of the cervical, thoracic and lumbar region of control (*R26R*$^{EnRCdx1/+}$) and *R26R*$^{EnRCdx1/+}$::*T-Cre*$^{Tg/+}$ (R26(TCre)$^{EnRCdx1/+}$) offspring at postnatal (P) day 1. (A) Snapshot of the cervical region. Note the presence of anterior transformations, malformed neural arches and vertebra fusions across the entire mutant cervical region. C1 and its characteristic anterior arch of the atlas (AAA) are fused to the basioccipital bone; C2 and C3 are anteriorly

transformed (denoted as "C1", "C2") and exhibit malformed neural arches; also note the absence of the tuberculum anterior (TA) in the mutant C6 vertebra as well as the shorter neural arch in C7. (B) Snapshot of the thoracic region; (B'': dorsal view) and (B'': ventral view showing the ribs attached to the sternum); thoracic vertebra (T) are identified by the presence of ribs. The thoracic vertebral region of some mutants is shorter (see B1), exhibiting 12 thoracic vertebras (T1-T12) instead of 13 (T1-T13). Fusion of ribs (red asterisk in B2) was also observed. (C) The lumbar region of the mutants is shorter presenting 5 lumbar vertebrae (L1-L5) instead of 6. (D) Summary of the skeletal phenotypes and frequency of axial skeletal abnormalities observed in n=6 R26(TCre)$^{EnRCdx1/+}$ mutant offspring analysed. AAA: anterior arch of the atlas; NA: neural arch; TA: tuberculum anterior; C: cervical vertebra; T: thoracic vertebra; L: lumbar vertebra.

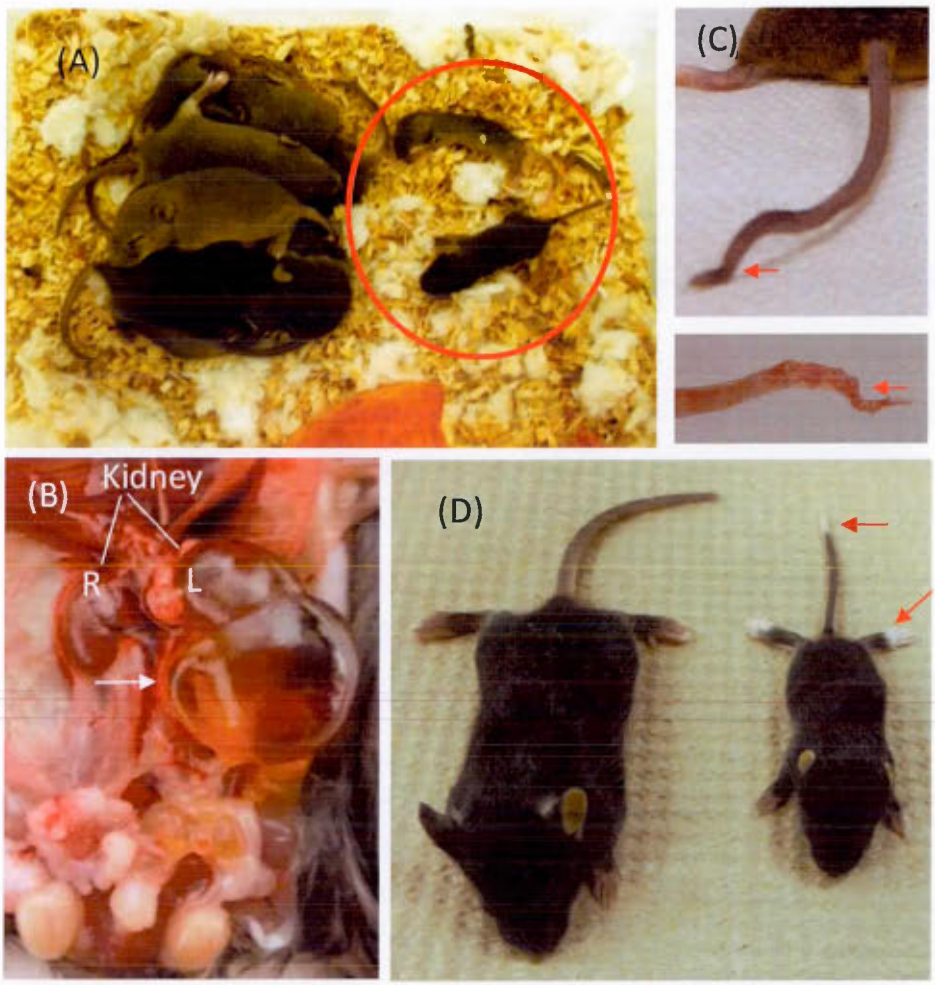


Figure 3.3 Phenotypes resulting from expression of EnRCdx1 in the neuroectoderm by using the *P3Pro*-Cre line.

R26R$^{EnRCdx1/EnRcdx1}$ mice were crossed with *P3pro*-Cre mice expressing Cre recombinase in the dorsal neural tube and neural crest cells under the control of the 1.6kb *Pax3* proximal promoter(Li et al., 2000). (A) Picture of a postnatal day (P) 10 litter containing two R26(P3Cre)$^{EnRCdx1/+}$ mice (in red circle), which are smaller than their littermates and died. (B) Urogenital system of a P90 R26(P3Cre)$^{EnRCdx1/+}$ mice. Note that the left (L) kidney is hydronephrotic and severely swollen in size, also note the dilatation of the renal pelvis (white arrow). The right (R) kidney appears normal. (C) Mutant mice also exhibit a kinked tail with malformed sacral vertebra (red arrows), shown is a P90 mouse. (D) Comparison between P20 wild type and R26(P3Cre)$^{EnRCdx1/+}$ littermates . Note that mutant mice are smaller, exhibit

posterior pigmentary anomalies, particularly lack pigmentation of the distal tail and hind paws (denoted in red arrows).

Table 3.1 Phenotypes of R26(P3Cre)$^{EnRCdx1/+}$ double transgenic mice

No. Matings	Number/Total (%)	Pigmentation defects (%)	Hydronephrosis (%)*	Mega-bladder*	Kinked tail (%)*
7	(14%) (19/139)	100% (19/19)	53% (8/15)	13% (2/15)	44% (4/9)

*Random group of R26(P3Cre)$^{EnRCdx1/+}$ mice, takes into account mice from other matings

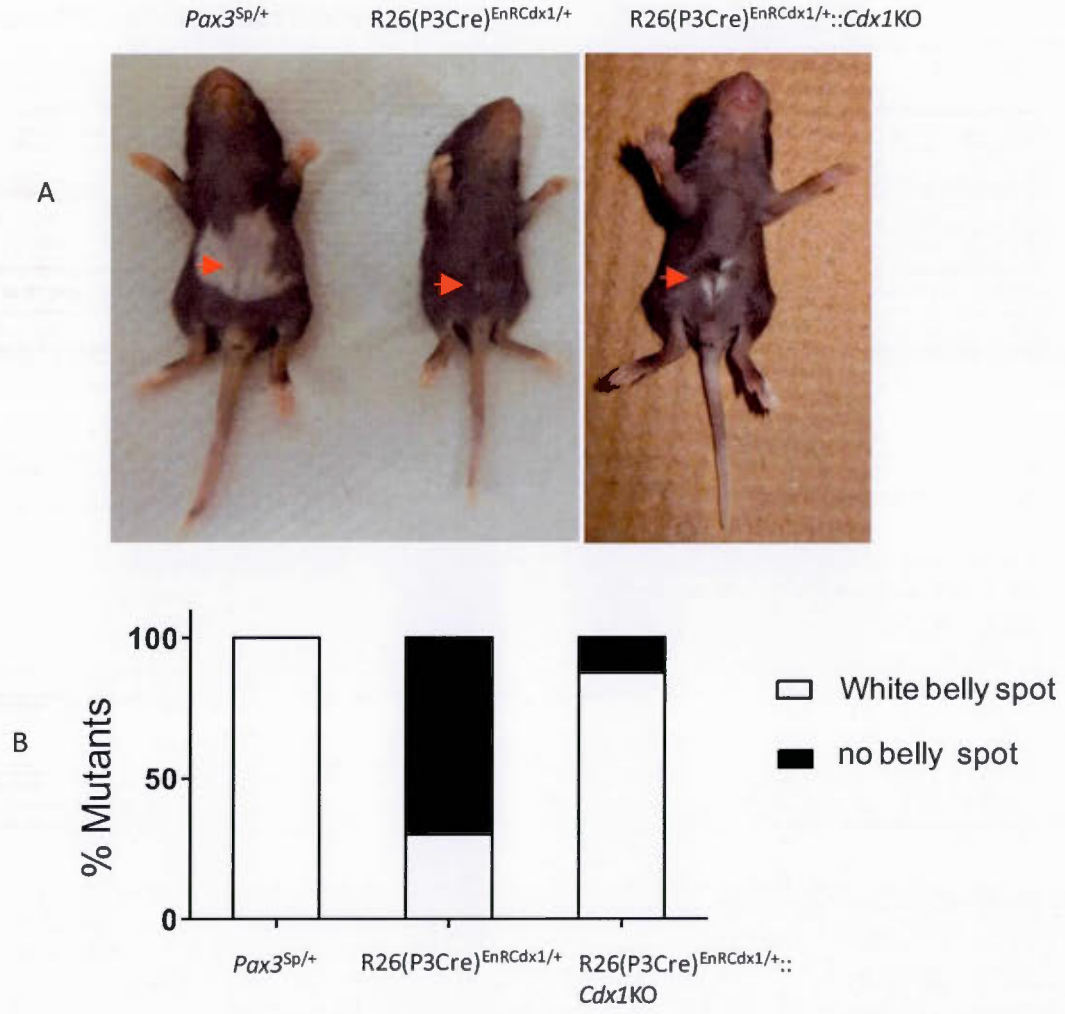


Figure 3.4 Cdx loss of function in the neuroectoderm recapitulates posterior pigmentary anomalies of *Pax3* Splotch ($Pax3^{Sp/+}$) mutants.

(A) Comparison of phenotypes between heterozygous $Pax3^{Sp/+}$, $R26(P3Cre)^{EnRCdx1/+}$ and $R26(P3Cre)^{EnRCdx1/+}$::*Cdx1*KO mutant mice at postnatal day (P) 10. $Pax3^{Sp/+}$ mice exhibit a characteristic white belly spot (red arrow), among other pigmentary anomalies. $R26(P3Cre)^{EnRCdx1/+}$ mice exhibit a very small white spotting in the belly (red arrow). Note that the size of the white spot increases in the *Cdx1* null background ($R26(P3Cre)^{EnRCdx1/+}$::*Cdx1*KO, red arrow) . (B) Percentage of mice that present a white belly spot in each of the three aforementioned genotypes. The penetrance of the white belly spot in

R26(P3Cre)$^{EnRCdx1/+}$ mutants increases from 40% to 87% when *Cdx* dosage is decreased (*Cdx1* KO background).

Table 3.2 Mendelian ratios of *R26R*$^{EnRCdx1/+}$, *R26R*$^{EnRCdx1/+}$:: *Pax3*$^{Sp/+}$, R26(P3Cre)$^{EnRCdx1/+}$, R26(P3Cre)$^{EnRCdx1/+}$::*Pax3*$^{Sp/+}$ allelic combinations

Genotype	*R26R*$^{EnRCdx1/+}$	R26R$^{EnRCdx1/+}$:: *Pax3*$^{Sp/+}$	R26(P3Cre)$^{EnRCdx1/+}$	R26(P3Cre)$^{EnRCdx1/+}$:: *Pax3*$^{Sp/+}$
Number/ Total (%)	30/63 (48%)	23/63 (37%)	6/63 (9%)	4/63 (6%)
Expected	25%	25%	25%	25%

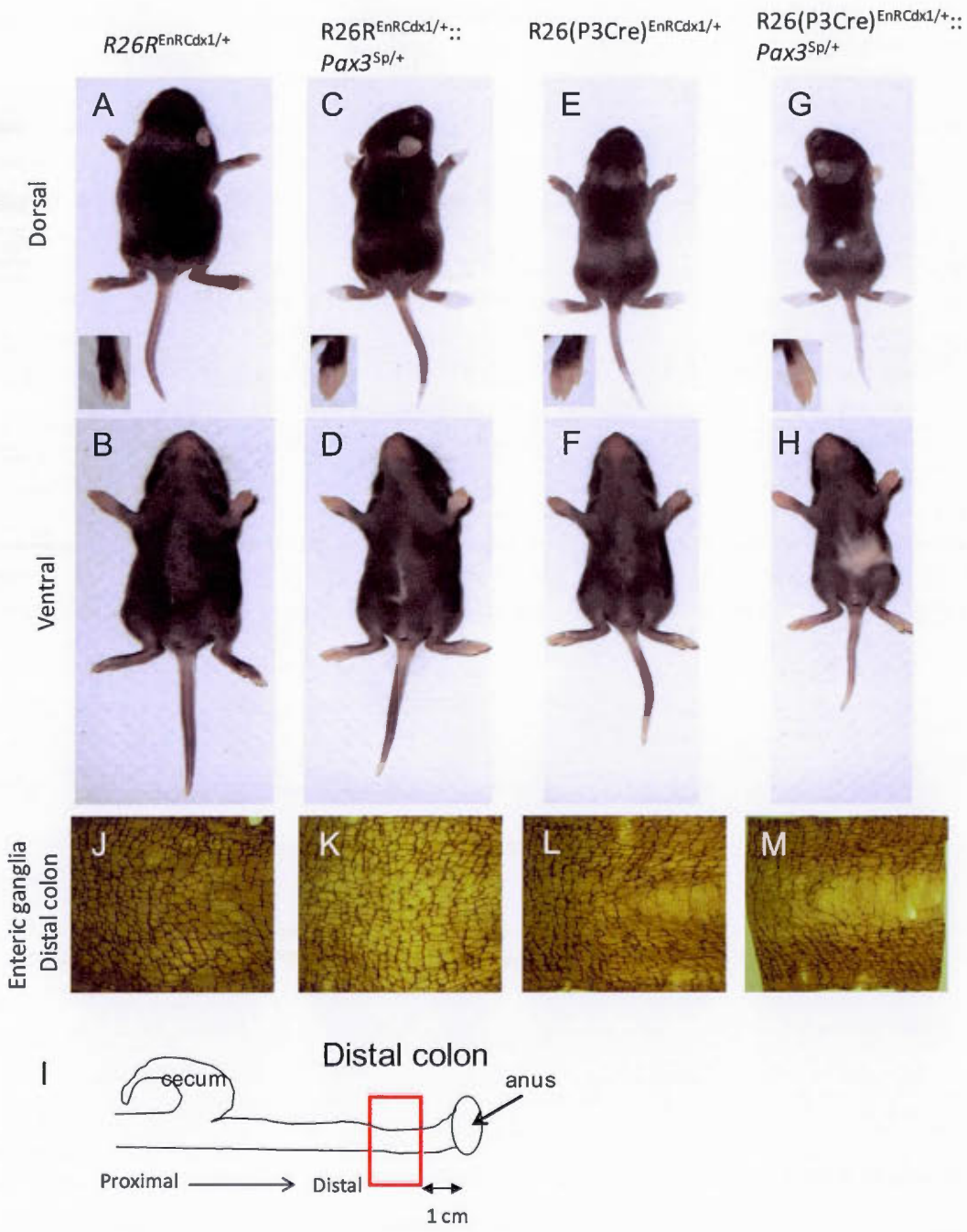


Figure 3.5 Cdx and Pax3 genetically interact in the control of melanocyte and enteric ganglia development.

(A-H) Dorsal (A-G) and ventral (B-H) views of control ($R26R^{EnRCdx1/+}$), (A,B); $R26R^{EnRCdx1/+}$::$Pax3^{Sp/+}$ (C,D); $R26(P3Cre)^{EnRCdx1/+}$ (E,F) and $R26(P3Cre)^{EnRCdx1/+}$::$Pax3^{Sp/+}$ (G,H) littermates at postnatal (P) day 10. Compared to Ctl animals (A,B), $R26R^{EnRCdx1/+}$::$Pax3^{Sp/+}$ mutants (C,D) lack pigmentation in the distal tail, fore and hind paws as well as present a white belly spot. $R26(P3Cre)^{EnRCdx1/+}$ mice (E,F) phenocopy the pigmentation defects of $R26R^{EnRCdx1/+}$::$Pax3^{Sp/+}$ mutants but only at posterior levels; note that the fore paws are pigmented and the white belly spot is barely visible. Posterior pigmentary anomalies of $R26(P3Cre)^{EnRCdx1/+}$ mice considerably increase in a $Pax3^{Sp/+}$ background (G,H). Note the more severe de-pigmentation of the tail, hind paws and belly and emergence of a white spot in the back of $R26(P3Cre)^{EnRCdx1/+}$::$Pax3^{Sp/+}$ animals that is not observed in mice of the other genotypes. n=3 independent observations in each genotype. (I-M) Acetyl cholinesterase staining in the distal colon region of P33 mice of the genotypes aforementioned described. (I) Schematic drawing of the colon region of the gut, showing the position of the analysed zone (red box) respect to the proximal (P)-distal (D) axis. Note the reduced density of the enteric nervous system in the distal colon of $R26(P3Cre)^{EnRCdx1/+}$ mice (L) and lack of enteric ganglia in the distal region compared to control (J). In $R26(P3Cre)^{EnRCdx1/+}$::$Pax3^{Sp/+}$ mice (M), the aganglionic zone extends across the entire analysed region. Acetyl cholinesterase staining results are representative of n=2 (for $R26R^{EnRCdx1/+}$ and $R26R^{EnRCdx1/+}$::$Pax3^{Sp/+}$) and n=1 ($R26(P3Cre)^{EnRCdx1/+}$ and $R26(P3Cre)^{EnRCdx1/+}$::$Pax3^{Sp/+}$) independent experiments.

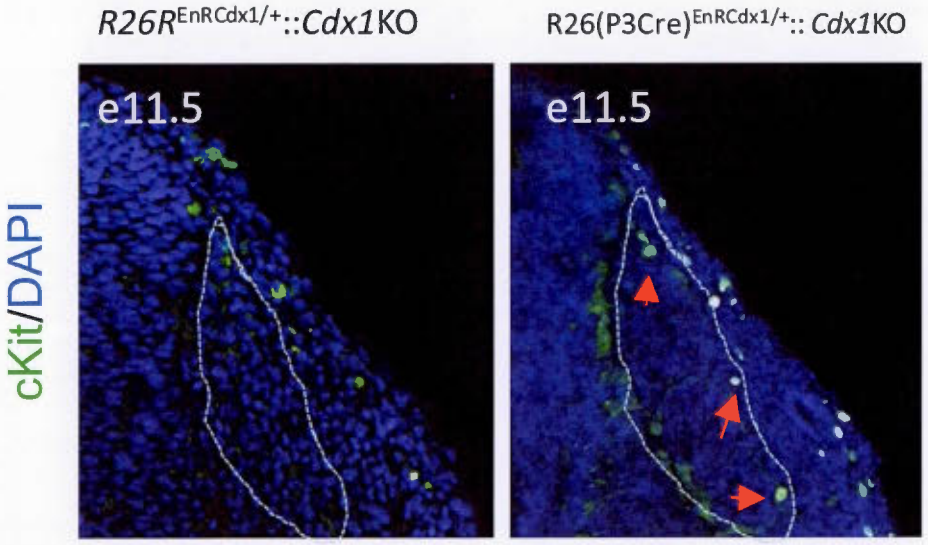


Figure 3.6 Analysis of the melanoblast population in R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO e11.5 embryos.

Shown is the analysis for expression of the melanoblast marker cKit in control (*R26R*$^{EnRCdx1/+}$ :: *Cdx1KO*) and R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO e11.5 embryos Note that the number of cKit positive melanoblasts is increased in mutant *R26R*$^{EnRCdx1/+}$: : *P3*pro-Cre$^{Tg/+}$:: *Cdx1* e11.5 embryos, n=2 independent observations; also note the presence of ectopic c-Kit positive cells in the dorsal root ganglia (red arrows).

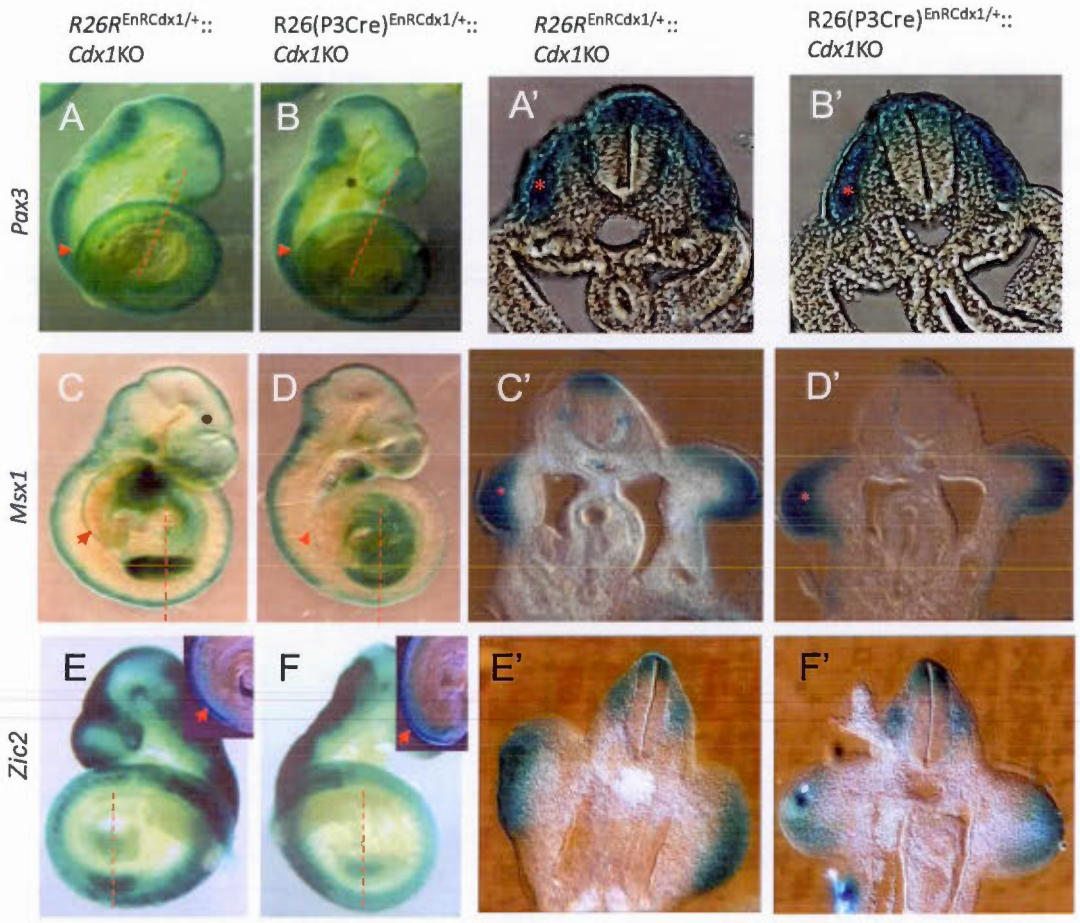


Figure 3.7 Cdx act early on the trunk NC-GRN by regulating posterior expression of the neural plate border specifiers *Pax3* and *Msx1*.

(A-F) Lateral views of Ctl (*R26R*$^{EnRCdx1/+}$ :: *Cdx1KO*), (A,C,E) and R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO (B,D,F) stage matched e9.5 embryos showing the expression of the neural plate border specifiers *Pax3 (A,B)*, *Msx1* (C,D) and *Zic2* (E,F), as evaluated by whole mount in situ hybridisation. The red dashed lines in (A-F) indicate the level at which the transverse sections shown in (A'-F') were made. Note that compared to Ctl, in R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO embryos *Pax3(B)* and Msx1(D) transcripts levels are specifically reduced in pre-migratory NCC (arrowhead) and the dorsal NT (B' and D') at the caudal region. Note that *Pax3* expression in the somite dermamyotome (red asterisk B') or *Msx1* expression in the limb bud (red asterisk D') remains unchanged in. Expression of *Zic2* is not affected in pre-migratory neural crest cells (arrowhead and inset in F) or dorsal neural tube (F') of

R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO embryos compared to Ctl (E,E'). Results are representative of n=3 embryos per ISH.

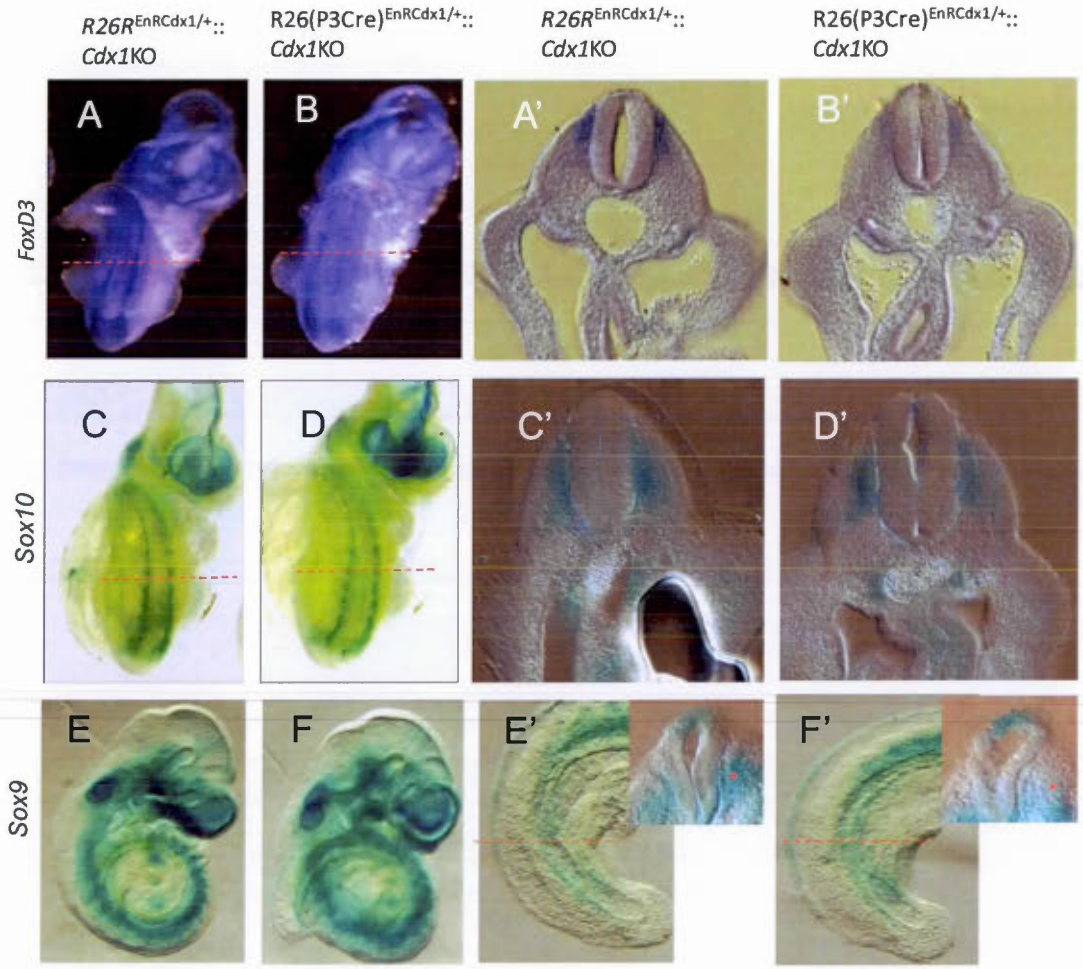


Figure 3.8 Expression of the neural crest specifiers *FoxD3* and Sox9 is deregulated in R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO e9.5 embryos.

(A-F) In situ hybridisation analysis of *FoxD3* (A,B,A',B'), *Sox10* (C,D,C',D') and *Sox9* (E,F,E',F') expression in Ctl (*R26R*$^{EnRCdx1/+}$ :: *Cdx1KO*), (A,C,E) and R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO (B,D,F) stage matched e9.5 embryos. The red dashed lines indicate the level at which the transverse sections shown in (A'-F') were made. Note the reduction in *FoxD3* expression in migratory neural crest cells in R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO mutants (compare A,B; A',B'), and the upregulation of *Sox9* transcripts in premigratory trunk neural crest cells (compare E',F') whereas *Sox10* expression remains unaffected. Results are representative of n=3 embryos per ISH.

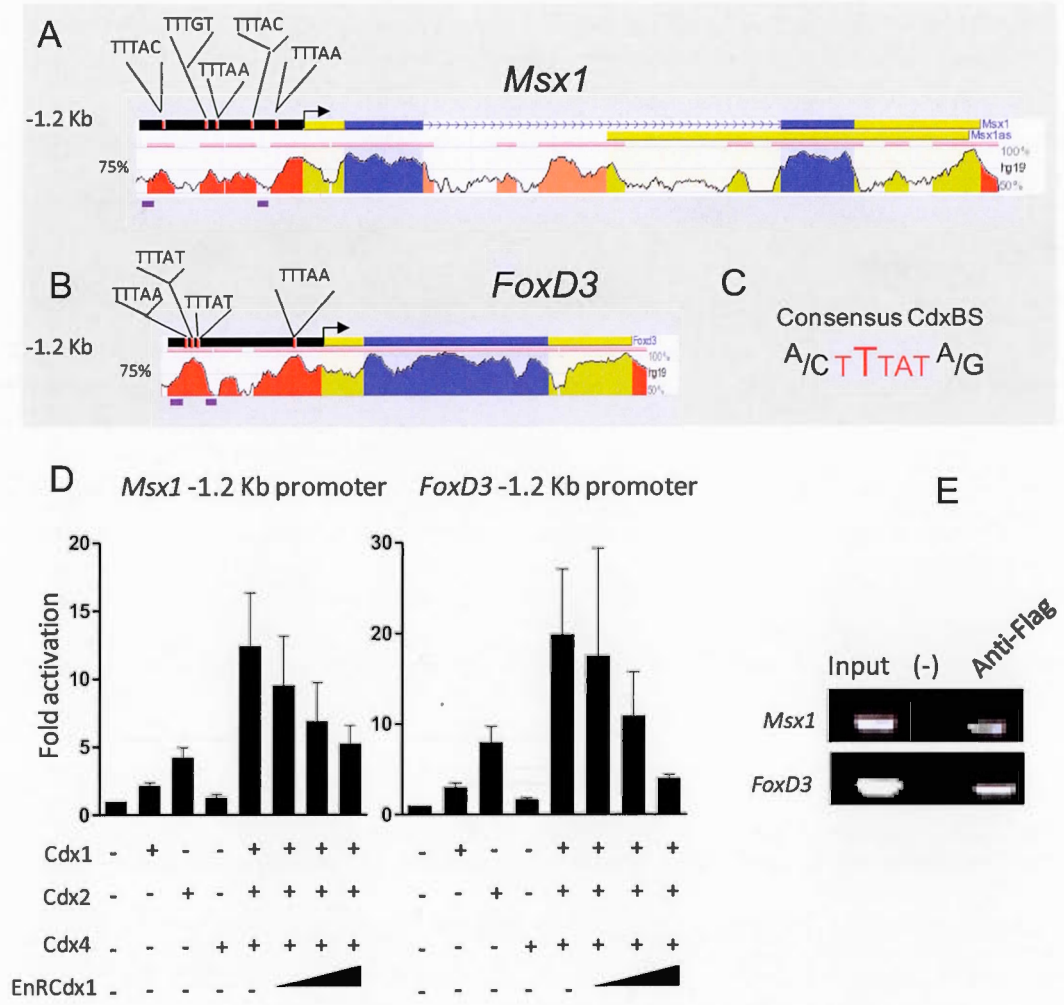


Figure 3.9 Cdx proteins regulate *Msx1* and *FoxD3* 1.2 kb proximal promoters.

(A,B) Comparative analysis of *Msx1* (A) and *FoxD3* (B) genomic regions using the ECR browser reveals the presence of several conserved sequences (denoted as red peaks and exhibiting 75% or more level of conservation between mouse and human) within the 1.2 kb proximal promoter region of both genes. See the enrichment in Cdx binding sites (CdxBS) in these regions. The consensus CdxBS is represented in (C). (D) Co-transfection assays in N2a cells using Luciferase reporter constructs driven by the 1.2 kb upstream proximal promoter sequences of *Msx1* and *FoxD3,* as well as Cdx1, Cdx2, Cdx4 and EnRCdx1 expression vectors. The results are expressed as fold induction compared to the relevant reporter vector alone. n=4 independent experiments performed in triplicate. Error bars indicate s.e.m. Note that Cdx1, Cdx2 and Cdx4 synergistically transactivate both proximal promoters and that co-

expression of EnRCdx1 represses in a dose-dependent manner the Cdx synergistic effect. (E) Anti-FLAG chromatin immunoprecipitation analysis from the trunk region of R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO e9.5 embryos shows occupancy by $_{FLAG}$EnRCdx1 of the *Msx1* and *FoxD3* proximal promoters in vivo. The genomic position of the primers used to amplify Msx1 and FoxD3 conserved sequences containing CdxBS is denoted as purple boxes in A and B, respectively. Note that these sequences are amplified from chromatin samples immunoprecipitated with anti-FLAG but not with the non specific normal rabbit IgG.

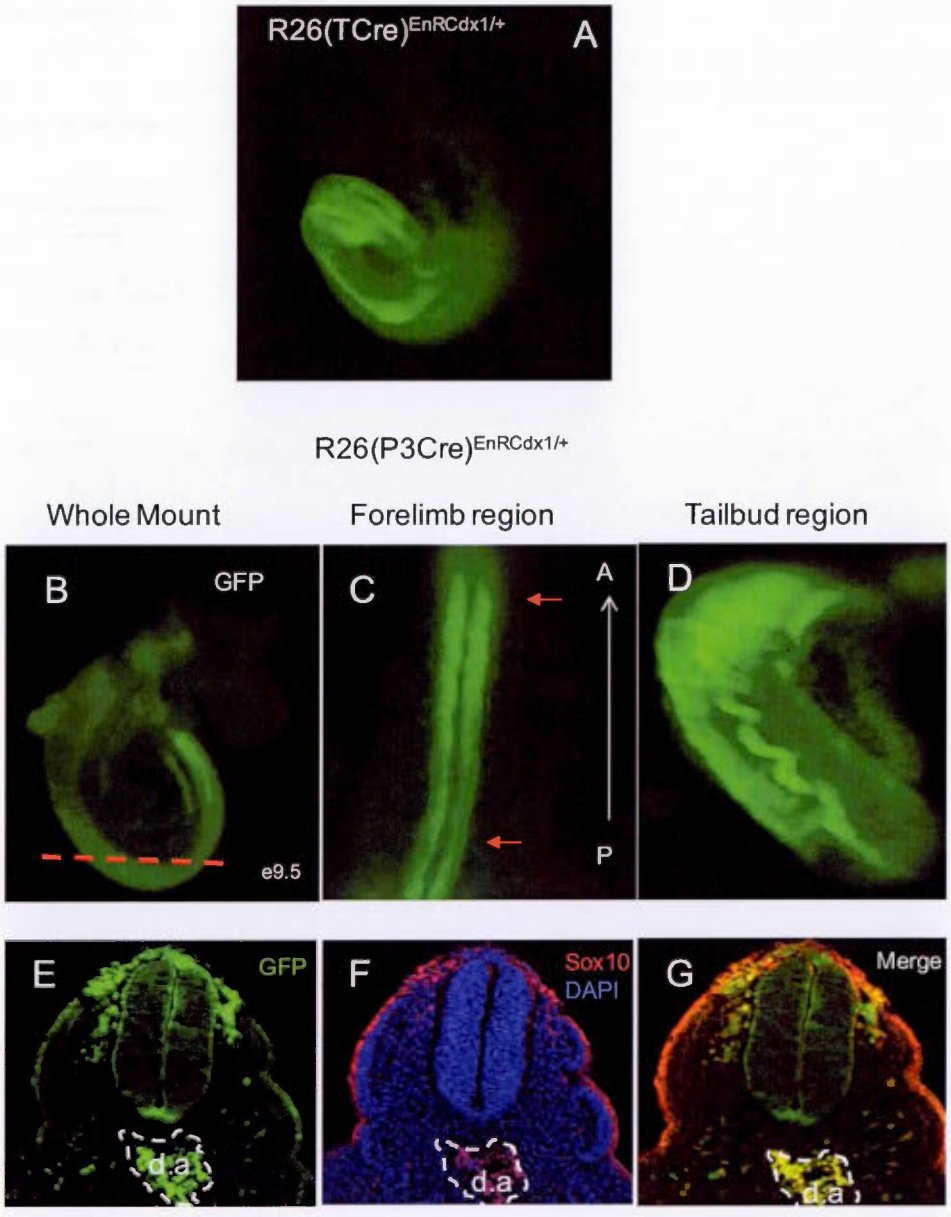


Figure 3.10 Validation of tissue-specific Cre mediated recombination and expression of EnRCdx1-IRES-EGFP in e9.5 embryos.

(A) Fluorescence (EGFP) image of a R26(TCre)$^{EnRCdx1/+}$embryo showing *T*-Cre mediated activation of eGFP in the posterior nascent mesoderm and derivatives. (B) Fluorescence (eGFP) image of a R26(P3Cre)$^{EnRCdx1/+}$ embryo (lateral view) showing *P3pro*-Cre mediated

activation of EGFP in the dorsal NT across the whole caudal axis from the hindbrain to the tailbud. The dashed red line represent the level at which the sections shown in (E,F,G) were made. (C) Dorsal view of the embryo at the forelimb region showing EGFP fluorescence specific to the NT and delaminating NCC. (D) EGFP fluorescence is broadly detected in the tailbud. (E,F,G) Vibratome transverse section at the level depicted by the dashed red line in B. (E) Note the strong EGFP fluorescence in migrating NCC, also note weak EGFP expression across the entire axis of the NT. (F) Immunofluorescence analysis for the NC marker Sox10 in the R26(P3Cre)$^{EnRCdx1/+}$ embryo. (G) Merged EGFP fluorescence and Sox10 immunostaining showing coexpression of GFP and Sox10 in NCC.

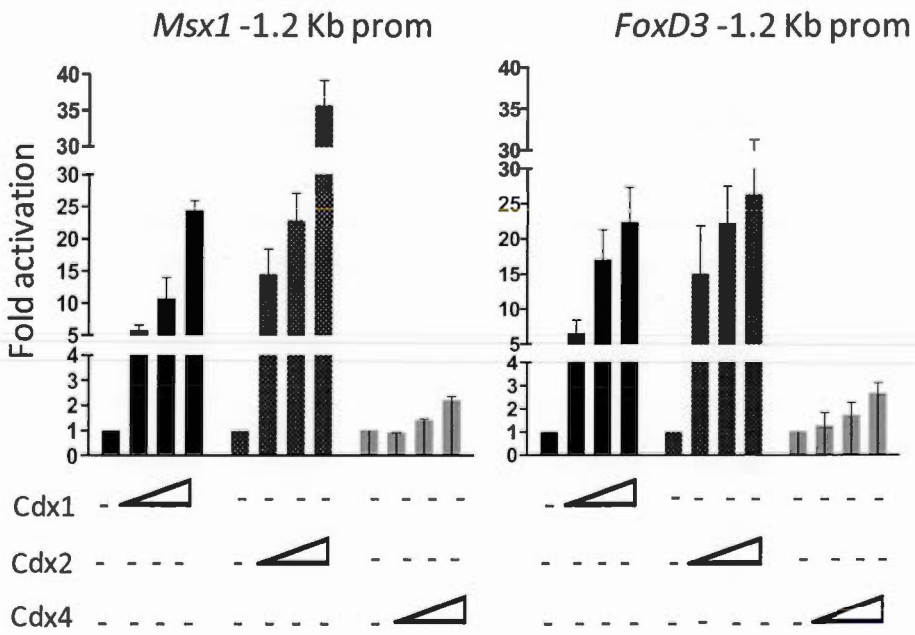


Figure 3.11 Cdx1, Cdx2 and Cdx4 differentially transactivate the Msx1 and FoxD3 1.2 kb proximal promoters.

Co-transfection assays in N2a cells using Luciferase reporter constructs driven by the 1.2 kb upstream proximal promoter sequences of *Msx1* and *FoxD3* as well as Cdx1, Cdx2, Cdx4 expression vectors. The results are expressed as fold induction compared to the relevant reporter vector alone. n=3 independent experiments performed in triplicate. Error bars indicate s.e.m. Note that addition of Cdx1, Cdx2 or Cdx4 expression construct transactivate in a dose-dependent manner Msx1 and FoxD3 promoter sequences. Also note that Cdx proteins exhibit differential transcriptional activity, with Cdx2 eliciting the strongest response, followed by Cdx1 and then by Cdx4.

Table 3.3 Mendelian ratios of R26(P3Cre)$^{EnRCdx1/+}$ and R26(P3Cre)$^{EnRCdx1/EnRCdx1}$ mice in WT, *Cdx1 KO* and *Pax3 Splotch* backgrounds.

Background	Genotype	No. Matings	Number/ Total	% / Expected
WT	R26(P3Cre)$^{EnRCdx1/+}$	7	19/139	14/50
	R26(P3Cre)$^{EnRCdx1/EnRCdx1}$	15	0/22	0/50
***Cdx1* KO**	R26(P3Cre)$^{EnRCdx1/+}$:: *Cdx1*KO	27	18/142	12/50
	R26(P3Cre)$^{EnRCdx1/EnRCdx1}$:: *Cdx1*KO	3	0/0	0/25
Splotch	R26(P3Cre)$^{EnRCdx1/+}$:: *Pax3*$^{Sp/+}$	10	4/63	6/25
	R26(P3Cre)$^{EnRCdx1/EnRCdx1}$:: *Pax3*$^{Sp/+}$	3	0/0	0/12.5

CHAPTER IV

General discussion

4.1 Pax3, the first Cdx target in the trunk NC-GRN: mechanistic studies

Our first study published in JBC in 2012 (Sanchez-Ferras et al., 2012) supported the general idea that Cdx proteins occupy a strategic position in the murine trunk NC-GRN downstream of the canonical Wnt input and more specifically resulted in a better understanding of the mechanisms that initiate *Pax3* neural expression in the caudal neuroectoderm. Indeed, on one hand, we demonstrated that Cdx proteins act as intermediates for relaying the canonical Wnt input to the NCE2 CRM while, on the other hand, we identified NCE2 as the minimal regulatory region able to recapitulate both the induction and the dorsal restriction of *Pax3* expression in the caudal neurectoderm. However, comparison of the wide Cdx protein distribution in the neuroectoderm with the delimited *Pax3* expression patterns indicated that more restricted inputs must cooperate with the Wnt-Cdx pathway to establish and refine *Pax3* expression in the NC domain. Such positional information could be provided either by a positive input from a Cdx cofactor specific for the neural plate border and dorsal NT or by a *Pax3* repressor specific for the medial neural plate and ventral NT, or alternatively by a combination of both mechanisms.

In the second study, published in BBA-GRM in 2014 (Sanchez-Ferras et al., 2014), we thus improved the understanding of the mechanism by which Cdx regulates *Pax3* expression and analysed how A-P positional information from the Wnt-Cdx pathway are co-integrated with D-V patterning cues onto NCE2. This work identified Zic2 as a novel direct regulator of *Pax3* NCE2 and Cdx neural cofactor. Our results led us to propose a model in which the Wnt-Cdx pathway is integrated with a Zic2 positive input to induce *Pax3* expression in the PNP and subsequently with a D-V positional cue from a Shh-Nkx6.1-Zic2 circuit to dorsally restrict *Pax3* expression in the closed NT. This model also includes a positive input from the SoxB1 member Sox2, which

may act in both the induction and dorsal restriction steps by providing neural specificity. However, neither Zic2 nor Sox2 expression is restricted to the lateral borders of the PNP, indicating that other factors must also cooperate with the Wnt-Cdx pathway.

Although a Shh-Nkx6.1 repressive input cannot account for the initial lateral restriction of *Pax3* expression in the open PNP (Sanchez-Ferras et al., 2014), we cannot rule out the possibility that another repressor may be involved in this regulation via the NCE2. And, according to its expression pattern in the tailbud region and known role as a repressor (Gofflot et al., 1997), the even skipped homeobox family member Evx1 would represent a good candidate. Nonetheless, our best candidate so far as being involved in *Pax3* lateral restriction via NCE2 is the canonical BMP-Smad signaling pathway. Indeed, high levels of BMP signaling activity are present in the E8.5 mouse neural plate border (Yang and Klingensmith, 2006), and the BMP input has being implicated in positioning of the neural plate border and induction of *Pax3* expression in chick and zebrafish, respectively (Garnett et al., 2012; Liem et al., 1995). Very interestingly, according to a MatInspector analysis of NCE2 sequences, two putative Smad binding sites are found in *Pax3* NCE2 in the vicinity of the Cdx and Zic binding sites and close to a putative Sox binding site (Figure 4.1). It is also noteworthy that Cdx proteins have been shown to physically interact with pSmad1/5/8 (Mari et al., 2014), Zic2 (Sanchez-Ferras et al., 2014) and Sox2 (Beland et al., 2004). Taken together, all these elements strongly suggest that pSmad1/5/8 might form a transcriptional complex with Cdx, Zic2 and Sox2 that turns on NCE2 activity and provides positional information to establish the lateral expression domain of *Pax3* in the caudal neuroectoderm.

We have already began to test this hypothesis by performing co-transfection assays in N2a cells and observed a very strong synergistic activation of a NCE2-luciferase reporter construct when co-expressing Cdx1, Zic2, Smad5 and Sox2 (Figure 1). This

study is currently the project of a Master's student in Dr. Pilon's lab and will be developed by performing the same battery of experiments as shown in Chapter I and II (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012) as well as by perturbation analysis of BMP signaling in the N2a cell line and *in vivo*. This study is expected to be the first to demonstrate how BMP and Wnt signaling pathways are integrated onto a single *Pax3* CRM, and finally provide the missing piece of the NCE2 regulatory circuit.

By characterizing the mechanisms of regulation of *Pax3* neural expression via the NCE2 in the mouse, we have tackled one of the central questions in developmental biology which consists of deciphering how positional information from A-P and D-V signaling cues is integrated at the *cis*-regulatory level to establish and refine gene expression patterns. Our study clearly shows that *Pax3*NCE2 acts as a "super-enhancer" that contains enough cis-regulatory information to integrate A-P and D-V cues and drive both induction and dorsal restriction of *Pax3* expression in the mouse caudal neurectoderm. However, this work only represents a small contribution to the understanding of the mechanisms that regulate expression of this important developmental gene. Our work and the work of others strengthen the idea that regulation of *Pax3* expression is complex and involves several redundant CRMs that respond to canonical Wnt, FGF, RA, BMP and Shh positional cues in a complementary manner. Several DNA binding factors have been shown to bind and transactivate these elements, including members of the Lef/Tcf, Tead, Pbx, Hox, Pou, Nkx, SoxB and Pax families (Degenhardt et al., 2010; Garnett et al., 2012; Milewski et al., 2004; Moore et al., 2013; Pruitt et al., 2004; Sanchez-Ferras et al., 2012; Zhao et al., 2013). We have extended this list by showing that Cdx1, 2 and 4 as well as Zic2 are novel direct regulators of *Pax3* neural expression (Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012).

Our studies in the mouse stress mechanistic differences in the interpretation of the *Pax3* regulatory circuit between species, as we have challenged the work of others suggesting that a Shh-Nkx6.1 circuit is involved in initial lateral restriction of *Pax3* expression at the PNP borders in less complex vertebrate species (Moore et al., 2013; Sanchez-Ferras et al., 2014). On the other hand, while we have evidenced a Wnt-Cdx-Pax3 pathway acting via NCE2 in the mouse, others have evidenced direct regulation by the canonical Wnt pathway via other CRMs in both the mouse and the zebrafish (Degenhardt et al., 2010; Moore et al., 2013; Zhao et al., 2014). As proposed in Chapter II (Sanchez-Ferras et al., 2014), it is probable that evolution created a Wnt-Cdx indirect and redundant regulatory circuit to buffer and preserve the expression of *Pax3* from fluctuations in the signaling inputs. Indeed, in ß-Catenin conditional KO mutants, *Pax3* expression is reduced but not abrogated (Zhao et al., 2014), suggesting that other inductive cues ensure expression of this important gene. Posteriorizing FGF and RA signaling have been involved in induction of *Pax3* neural expression (Bang et al., 1999; Bang et al., 1997; Garnett et al., 2012) and Cdx are well known integrators of these signals (Lohnes, 2003). Deletion analysis of the 1.6 kb proximal promoter in P19 cells have shown that the region comprising NCE1 and NCE2 is responsive to RA signaling (Natoli et al., 1997). Therefore, it would be interesting to explore whether Cdx can also relay positional information from RA and FGF signaling pathways to the *Pax3* promoter and whether these inputs are integrated on the NCE2.

4.2 *Msx1* and *FoxD3* as two novel Cdx targets in the trunk NC-GRN

The goal of the EnRCdx1 mouse project was to identify putative targets of Cdx in the trunk NC-GRN. The generation of the EnRCdx1 mouse model was time consuming and took 2 years, but it was worth it. Combining this tool with the *P3Pro*-Cre line led us to knockdown Cdx function in the NC *in vivo* and to identify *Msx1* and *FoxD3* as two novel Cdx targets. Both genes are downstream of canonical Wnt signaling (Bang et al., 1999; Sato et al., 2005), suggesting that Cdx serve to relay this input to the

Msx1 and *FoxD3* proximal promoter as in the case of *Pax3*. As we did for *Pax3*, future work will be required for characterizing the mechanism by which Cdx proteins regulate the expression of these important NC regulators. The presence of clustered CdxBS in the 1.2 kb proximal promoter of *Msx1* and *FoxD3*, and our ChIP data showing occupancy of these conserved regions by EnRCdx1 *in vivo* (also supported by Cdx2 ChIP-seq data from Dr. Lohnes' lab), indicate that Cdx proteins directly regulate *Msx1* and *FoxD3* expression via CdxBS. However, EMSA analysis will be needed to confirm direct Cdx binding to these elements. According to our sequence analysis, both promoters also contain putative Lef/Tcf binding sites, raising the possibility that Cdx proteins might act independently of DNA binding by interacting with the ß-Catenin-LEF/TCF complex via LEF/TCF response elements (Beland et al., 2004).

It is also important to bear in mind that the wide expression of Cdx proteins in the PNP and caudal NT implies that these transcription factors must cooperate with other factors to restrict expression of *Msx1* and *FoxD3* to the NC. As we did for *Pax3* (Sanchez-Ferras et al., 2014), future projects must analyze this mechanism. To the best of our knowledge, a study of the regulatory regions that participate in establishment and refinement of *FoxD3* and *Msx1* expression in the trunk NC has not been performed in the mouse yet. It will first be necessary to test whether the Cdx-responsive proximal promoter of *Msx1* and *FoxD3* exhibit trunk NC activity or whether they are implicated in general activation of gene expression. Transgenic reporter studies in E8.5-E9.5 mouse embryos should provide the response to this. If these regions present NC activity, it will be necessary to study the mechanisms by which the Cdx input is made specific for the NC domain. In the case of *Msx1,* as mentioned in Chapter 4, interesting candidates as NC-specific Cdx cofactors are the downstream effectors of BMP signaling pSmad1/5/8. In the case of *FoxD3*, studies in *Xenopus* and chick have identified Zic proteins, and particularly Zic1, as critical direct regulators of *FoxD3* expression in trunk NCC (Monsoro-Burq et al., 2005; Sato

et al., 2005; Simoes-Costa et al., 2012). This is particularly interesting since we have recently shown that Cdx proteins functionally and physically interact with Zic2 (Sanchez-Ferras et al., 2014) and Zic members are known to exhibit functional overlap and recognize similar targets (Merzdorf, 2007). The anterior-specific expression pattern of Zic1 in E9.5 mouse embryos is not consistent with this gene being involved in induction of *FoxD3* expression in trunk NCC (see (Merzdorf, 2007). The Zic1 role in NCC development and *FoxD3* induction is most likely fulfilled by its redundant paralog Zic2 at more posterior levels. Indeed, *Zic2* Kumba mutants exhibit trunk NC defects and severe reduction in *FoxD3* expression (Elms et al., 2003). This raises the possibility that a Cdx-Zic2 interaction is at work on the proximal promoter region of *FoxD3* in order to restrict the Cdx input to the NC domain. This could simply be evaluated by performing co-transfection assays in N2a cells and testing whether Zic2 and Cdx1 synergistically transactivate the *FoxD3* 1.2 kb proximal promoter, followed by EMSA, ChIP and Co-IP analysis as described in Chapter II (Sanchez-Ferras et al., 2014). *In vivo*, this could be evaluated by generating a transgenic line that expresses Zic2 in the whole NT region and assess whether *FoxD3* expression is expanded ventrally.

If the *Msx1* or *FoxD3* proximal promoters do not exhibit NC activity, it is possible that Cdx proteins are required for the general activation of these regions in order to allow enough functional expression levels of *Msx1* and *FoxD3*. Indeed, as we and others have shown for several Cdx target genes (Charite et al., 1998; Sanchez-Ferras et al., 2012; Savory et al., 2011a; Subramanian et al., 1995), Cdx proteins appear to have a common *modus operandi* for the regulation of their targets via clustering on the proximal promoter regions. *Msx1* transcription requires tethering and stabilization of a multisubunit complex comprising TBP, Sp1 and the co-activator CBP/p300 on the minimal promoter situated -165/+106 bp relative to the TSS (Shetty et al., 1999). It has been proposed that Msx1 physically interacts with these proteins and squelches the transcription initiation complex in order to repress its own expression (Shetty et

al., 1999). Interestingly, there is a cluster of Cdx binding sites in close proximity to the Msx1 and Sp1 binding sites, and Cdx2 has been shown to physically interact with the co-activators p300 and CBP (Hussain and Habener, 1999; Lorentz et al., 1999). This suggests that Cdx proteins may be required to prevent Msx1 binding and thereby stabilize the transcription initiation complex by attracting and tethering transcriptional co-activators. Such a mechanism is used for Cdx2 in the activation of the *glucagon* promoter and other intestinal epithelial cell gene promoters (Hussain and Habener, 1999). Thus, it appears that Cdx proteins have affinity for "Mega-Trans"-containing enhancers, recently proposed as potent functional enhancers that recruit a MegaTrans complex of transcription factors, co-activators and transcription machinery (Liu et al., 2014).

We also cannot rule out the possibility that Cdx proteins regulate *Msx1* and *FoxD3* expression by operating via other CRMs. Deletion and transgenic reporter analysis in the mouse have identified a CRM named PE (MacKenzie et al., 1997). As previously described (Chapters 1 and 3), PE is 78 bp long and is located -2630 to -2553 bp upstream of the TSS. This small enhancer recapitulates *Msx1* expression in the roof plate along the entire A-P axis at E10.5 and other tissues and requires the proximal promoter region for robust expression. It is unclear whether this element drives early induction at the PNP border, although it has the capacity to respond to canonical Wnt and BMP signaling pathways via functional Tcf4 and Smad binding sites, respectively (Alvarez Martinez et al., 2002; Miller et al., 2007). According to our sequence analysis, at least one putative CdxBS is present in close proximity to the Tcf4 and Smad binding sites of the PE, suggesting that Cdx proteins may regulate this element via direct binding or via interaction with pSmad1/5/8 and/or the Lef/Tcf-ß-Catenin complex (Beland et al., 2004; Mari et al., 2014). Regarding *FoxD3*, studies in the chick have uncovered a trunk-specific NC enhancer (NC2) that directly responds to Msx1, Pax7, and Zic1 stimulation (Simoes-Costa et al., 2012). If this CRM is active in mice, it must be regulated by Pax3 and Zic2/Zic5 instead of Pax7

and Zic1 which are both not expressed in the most caudal neuroectoderm where NCC are newly induced at E9.5 (Mansouri et al., 1996; Merzdorf, 2007). Based on our recent work, it is thus tempting to hypothesize that Cdx may again functionally cooperate with Zic2 for *FoxD3* induction via the putative NC2 (Sanchez-Ferras et al., 2014).

4.3 Cdx are at the head of the trunk NC-GRN in the mouse

Because of genetic redundancy and the difficulty to perform perturbation analyses in the mouse, it is not well understood how the NC-GRN is wired in this higher vertebrate species. Our work implicating *Cdx* genes in the mouse NC-GRN and their known restricted posterior expression in the embryo, stresses that the trunk and cranial NC-GRNs are differentially wired. According to our results, *Cdx* genes occupy a strategic position at the head of the trunk NC-GRN downstream of the inductive inputs, and directly upstream of the neural plate border specifiers (Msx1, Pax3) as well as the NC-specifier module (FoxD3). Our results also suggest that Cdx proteins indirectly control *Sox9* transcript levels in pre-migratory NCC. We propose that FoxD3 may mediate this regulation by acting as a *Sox9* repressor, taking into account that *Sox9* expression in osteochondroprogenitors is precociously induced in NC-*FoxD3* cKO embryos (Mundell and Labosky, 2011). However, future work will be needed to confirm this hypothesis since both *Sox9* and *FoxD3* are normally co-expressed in pre-migratory trunk NCC and epitasis analyses in chick have shown that Sox9 lies upstream of FoxD3 (Nitzan et al., 2013). Moreover, we have demonstrated that Cdx proteins can physically interact with other components of the NC-GRN such as the neural plate border specifier Zic2 and also with the pan-neural factor Sox2. As we described above, given that Zic proteins have been implicated in the regulation of *FoxD3* expression in trunk NCC and that Cdx physically interact with Zic2 (Sanchez-Ferras et al., 2014; Simoes-Costa et al., 2012), it is probable that a Cdx-Zic2-Sox2 transcriptional complex is also implicated in the regulation of *FoxD3* as for *Pax3*.

Our work in the mouse together with the work of others in *Xenopus* suggest that Cdx, Gbx2 and Ap2a occupy a similar hierarchical position at the head of the vertebrate NC-GRN (de Croze et al., 2011; Li et al., 2009; Sanchez-Ferras et al., 2014; Sanchez-Ferras et al., 2012). Like *Cdx* genes, *Gbx2* and *Ap2a* have been shown to act as intermediates between canonical Wnt signaling and the neural plate border specifiers *Pax3* and *Msx1* (de Croze et al., 2011; Li et al., 2009; Sanchez-Ferras et al., 2014). The relationship between Cdx and Gbx2 is particularly intriguing, as both posteriorizing factors interact with Zic members to regulate *Pax3* expression (Li et al., 2009; Sanchez-Ferras et al., 2014). However, in contrast to Cdx and AP2a which have been implicated as direct activators of *Pax3* expression (de Croze et al., 2011; Sanchez-Ferras et al., 2012), Gbx2 is a repressor and thus cannot induce *Pax3* expression by itself (Li et al., 2009). Although it remains to be demonstrated, it is probable that Gbx2 represses a *Pax3* and *Msx1* repressor in order to permit activation by AP2a and Cdx proteins. The putative Gbx2-AP2a and Gbx2-Cdx regulatory circuits are in place at the moment of trunk NC induction in the mouse, with AP2a expressed in the posterior neural folds (Mitchell et al., 1991) and the others widely expressed in the caudal embryo (Meyer and Gruss, 1993; Wassarman et al., 1997). *Gbx2* expression strongly decreases at e8.5, but is maintained in the tailbud (Wassarman et al., 1997). This is consistent with a putative role for Gbx2 in neural plate border specification but not in subsequent steps of trunk NC development. In anterior regions of the mouse embryo, Gbx2 and AP2a play critical roles in cardiac and cranial NC development, respectively (Byrd and Meyers, 2005; Schorle et al., 1996). As far as we known, no trunk NC defects have been described in *Gbx2*-null mouse embryos. On the other hand, knockdown of AP2a in the mouse NC affects pigmentation of the paws, tail and belly (Brewer et al., 2004), similar to the effects caused by EnRCdx1-mediated Cdx loss-of-function. However, in contrast of Cdx, AP2a loss-of-function in the mouse does not appear to affect *Pax3* expression (Schorle et al., 1996). Thus, whether the putative Gbx2-AP2a and Gbx2-Cdx circuits

are at work and controls trunk NC development in the mouse remains to be determined.

4.4 Novel Cdx role in the control of melanocyte development in vertebrates via *Pax3*, *FoxD3* and *Sox9*

The study described in Chapter 4 is the first implicating Cdx proteins in the control of pigment cell development in vertebrates. Pigment cells are the most ancestral NC lineage and have recently been implicated in the origin of the NC (Abitua et al., 2012; Jeffery, 2006; Jeffery et al., 2008; Jeffery et al., 2004). Very interestingly, consistent with our work, one study reported that loss-of-function of a Cdx homolog in the ascidian embryo *Halocynthia roretzi* also affected pigmentation (Katsuyama et al., 1999). Although ascidians do not possess *bona fide* migratory NCC, homologs of the vertebrate NC-GRN including canonical Wnt signaling *Pax3*, *Msx*, *Zic* and *FoxD* are already in place at the putative neural plate border region (Abitua et al., 2012; Lemaire et al., 2002). Based on our work showing the regulation of these core NC genes by the Wnt-Cdx pathway in the mouse, it is thus tempting to speculate that *Cdx* are ancestral components of a rudimentary trunk NC-GRN present in urochordates and that this function has been evolutionarily conserved in vertebrates. It would be interesting to perform a comparative study of Cdx functions in the rudimentary ascidian NC-GRN.

Our analyses in the mouse show two main defects related to melanocyte development when knocking down Cdx function in the NC via expression of EnRCdx1. We observe an increase in melanogenesis at embryonic stage (e11.5), particularly an increase in the number of melanoblasts migrating along the dorsolateral pathway, with some ectopic melanoblasts also detected in neural derivatives such as DRG. However, adult mice intriguingly exhibit hypopigmentation and reproduce the characteristic posterior pigmentary anomalies of $Pax3^{Sp/+}$ mutants. Our analysis at E9.5 suggests that loss of Cdx function affects the expression of critical regulators of

melanocyte specification *Pax3*, *FoxD3*, and *Sox9*. Embryonic phenotypes, including the increase in the number of melanoblasts and their ectopic presence in the DRG seem to recapitulate the phenotypes of *FoxD3* cKO e11.5 mouse embryos (Nitzan et al., 2013). In these mutants, ectopic melanoblasts in sensory ganglia is also accompanied by a decrease in expression of neural and glial markers, suggesting a neural to melanocytic shift, according to the known role of FoxD3 as repressor of melanogenesis (Nitzan et al., 2013). Consistent with this observation, preliminary observations suggest reduced expression of the glial marker S100-ß in sensory ganglia of the e11.5 EnRCdx1 mutant embryos (data not shown). On the other hand, postnatal hypopigmentation also appears to be a consequence of *Pax3* deregulation in the mutants since penetrance of EnRCdx1 phenotype increase in a $Pax3^{Sp/+}$ background. Of note, ectopic and persistent expression of Sox9 in melanoblasts from e11.5 onwards as seen in Ods heterozygous mouse mutants is similarly associated with hypopigmentation and ventral spotting (Qin et al., 2004). However, whole mount *in situ* hybridization analysis of *Sox9* expression at e9.5 suggest that the upregulation of *Sox9* expression in the mutants is transient and specific to pre-migratory NCC, as *Sox9* expression appear to be normally down-regulated in migratory NCC at more anterior levels. So far, we don't know to what extent the transient deregulation of *Sox9* at e9.5 may impact in the phenotypes observed at later stages. In chick embryos, *Sox9* overexpression in the dorsal NT at early stages leads to increased melanoblast differentiation and premature migration along the dorsolateral pathway suggesting that early upregulation of *Sox9* expression contributes to the increased melanogenesis observed in e11.5 EnRCdx1 mutants. However, a conclusion of this work must wait for expression analyses of melanoblast markers including the master melanocyte regulators Mitf-M and Dct in EnRCdx1 mutants. Consistent with a *FoxD3* phenotype, ectopic and increased number of Mitf+ melanoblast are expected in the mutants. Epistasis analysis should also be performed in order to determine the contribution of *Pax3*, *FoxD3* and *Sox9* deregulation to the

phenotypes observed. In this regard, it is noteworthy that we tried to perform a rescue of *Pax3* expression in *R26(P3Cre)*$^{EnRCdx1/+}$ mice by mating with the *R26R*$^{Pax3/Pax3}$ line (Wu et al., 2008) but failed to recover mice with both EnRCdx1 and Pax3 knockin alleles (*R26(P3Cre)*$^{EnRCdx1/Pax3}$).

In this work, we were only able to analyze the EnRCdx1 heterozygous phenotype due to the low breeding capacity of *R26(P3Cre)*$^{EnRCdx1/+}$ mice. It is therefore possible that other putative roles for Cdx in NC development might be masked by incomplete Cdx loss-of-function. Future projects will need to use other NC-specific Cre lines, for example a tamoxifen-inducible Wnt1-Cre line (Wnt1-CreER^{T} available at The Jackson Laboratory) in order to test whether we can circumvent the early postnatal lethality and severe urogenital defects observed with the P3Pro-Cre line and to complete the characterization of the NC defects in the EnRCdx1 homozygous background. According to our finding that *Pax3* and *Foxd3* are Cdx targets and their known essential roles in trunk NC development (Nelms et al., 2011; Serbedzija and McMahon, 1997; Teng et al., 2008), a recapitulation of the severe NC phenotypes of *Pax3*$^{Sp/Sp}$ and NC-*FoxD3* cKO mutants is expected in the EnRCdx1 homozygous background. These may include excessive NCC apoptosis and severe reduction in the pool of trunk NC progenitors, lack/reduction of peripheral and enteric nervous system probably leading to megacolon. Based on the observed *Sox9* upregulation in EnRCdx1 heterozygous mutants and considering that this has been shown to promote NC specification in chick embryos, it is also possible that an increase in *Sox9* expression might compensate for the loss of *Pax3* and *FoxD3* in NC specification in the EnRCdx1 homozygous background (Cheung and Briscoe, 2003). On the other hand, although Cdx generally act as activators of gene expression (only two negative targets being reported as far as we know: *Mafb* and the *insulin-like growth-factor-binding protein 3* gene, see Chapter 1), one limitation of the EnRCdx1 model is that putative direct negative targets of Cdx proteins cannot be identified via this approach. It would therefore be interesting to complement the work done with the EnRCdx1

model with other conditional approaches (i.e. NC-specific *Cdx1/2* cDKO mutants) in order to determine the full scope of Cdx NC functions and targets (Savory et al., 2011a).

4.5 Tissue-autonomous Cdx role in the control of NT closure

Consistent with other work showing the implication of Cdx in the control of NT development (Savory et al., 2011a; van de Ven et al., 2011), we occasionally observed NT malformations in *R26(P3Cre)*$^{EnRCdx1/+}$ embryos (Figure 2). These defects were present at low frequency (<5%) and with an apparent posterior to anterior gradient of severity, probably reflecting differential effectiveness of EnRCdx1-mediated Cdx loss-of-function along the A-P axis. Indeed, in the EnRCdx1 model, the efficacy of Cdx knockdown depends on endogenous Cdx protein levels as EnRCdx1 must compete with them for common target genes or interacting partners. Since Cdx proteins are distributed in a posterior (high) to anterior (low) gradient along the posterior axis, phenotypes are expected to be more severe at anterior levels and mainly in the hindbrain-spinal cord region, where of the three Cdx proteins only Cdx1 is expressed (Lohnes, 2003; Meyer and Gruss, 1993). This enterily depends on where and when Cdx proteins function. However, consistent with the former hypothesis, open NTDs were more frequently observed at more anterior levels of the spinal cord whereas, in more posterior regions, a closed but wavy and irregularly shaped NT was more frequently observed (Figure 4.2).

Severe open NT defects (craniorachischisis) affecting the closure site 1 near the hindbrain-spinal cord boundary were observed in *Cdx1/2* cDKO mutants, due to defective regulation of convergent-extension movements via reduced expression of the PCP signaling receptor *Ptk7* (Savory et al., 2011). On the other hand, milder neuroepithelium arrangement defects were described in the caudal NT of *Cdx2/4* cDKO mutants. In this region, Cdx have been involved in organization of the neuroepithelium and apico-basal polarity in a canonical Wnt and Hox-dependent

manner (van de Ven et al., 2011). Although different in nature, the phenotypes of EnRCdx1 mutants suggest that Cdx proteins may operate in NT closure in a tissue-autonomous manner.

Differences in NT phenotypes of EnRCdx1 mutants along the A-P axis are consistent with region-specific mechanisms for Cdx proteins in the control of NT development. In this regard, the Wnt-Cdx pathway is expected to impact on NT closure in the spinal cord region, via regulation of *Pax3*. *PAX3* loss-of-function in human and mice results in a posterior NT closure defect known as *spina bifida aperta* (Agopian et al., 2013; Epstein et al., 1991). Although the reduced *Pax3* expression levels in EnRCdx1 mutants (in WT, *Cdx1*KO or Splotch backgrounds) appear to be sufficient to control normal PNP closure, *spina bifida aperta* is expected in the EnRCdx1 homozygous background. On the other hand, low levels of *Pax3* expression have been associated with defective roof plate morphology in ß-catenin cKO embryos (Zhao et al., 2014). Pax3 is known to control apoptosis in the dorsal NT and to probably influence cell surface properties (Mansouri et al., 2001; Pani et al., 2002; Wiggan et al., 2002). Thus, we cannot rule out the possibility that defective *Pax3* expression in *R26(P3Cre)*$^{EnRCdx1/+}$ e9.5 embryos contribute to the NT morphology defects observed at more caudal levels. However, EnRCdx1 mutants rather exhibit irregular shaping of the lumen neuroepithelium along the whole D-V axis, suggesting that Cdx proteins may also control organization of this tissue via *Pax3*-independent mechanisms. Of note, the presence of NTDs did not impair NC induction or delamination in *R26(P3Cre)*$^{EnRCdx1/+}$ e9.5 embryos, according to immunofluorescence analysis for the NC marker Sox10 in cryosections at the level of the lesion (Figure 4.2F). This is consistent with the known functional dissociation between spinal cord neurulation and NC delamination (Estibeiro et al., 1993; Franz, 1992). The molecular bases of the NTDs observed in *R26(P3Cre)*$^{EnRCdx1/+}$ embryos will need to be studied in the future.

4.6 Novel Cdx function in the control of ureter development

Our characterization of *R26(P3Cre)*$^{EnRCdx1/+}$ animals has also unveiled an unexpected key Cdx role in ureter development and function as revealed by the observation of severe hydronephrosis (increase in the diameter of the kidney renal pelvis and collecting duct system) (Figure 4.3 A,B). Hydronephrosis is the most frequent congenital abnormality detected by prenatal ultrasound (1% incidence) and a common cause of renal failure in children (Chevalier, 1998; Dudley et al., 1997). In this serious disease, renal pelvis and kidney collecting duct system dilate due to the fluid pressure caused by urine accumulation, which progressively destroys the renal parenchyma, as we observed in the hydronephrotic kidney of *R26(P3Cre)*$^{EnRCdx1/+}$ mice (Figure 4.3C,C'). The molecular mechanisms implicated in this serious disease are largely unknown.

Hydronephrosis can be classified in two forms: obstructive or non-obstructive. Physical obstruction may be caused by defects in the functional integrity and cell layer organization of the renal pelvis or ureteric wall. Obstruction is frequently observed at the level of the ureteropelvic junction (Chang et al., 2004). Non-obstructive hydronephrosis is commonly associated with defects in ureter peristalsis, which is the unidirectional movement that propels the urine from the kidney down to the bladder (Bohnenpoll and Kispert, 2014; Cain et al., 2011). Ureter peristalsis and then urine evacuation is ensured by coordinated waves of contractions of the ureter smooth muscle cell (SMC) layer, propagating in a proximal to distal direction relative to the kidney. These waves of contractions are controlled by the spontaneous electrical activity of two different ureteric pacemaker populations of cells: one present in the pelvis-kidney-junction and marked by the expression of hyperpolarisation-activated cation-3 channel (Hcn3) (Hurtado et al., 2010) and the other one present throughout the ureter, known as interstitial cells of Cajal-like cells (ICC-LCs) and marked by Kit expression (David et al., 2005; Lang et al., 2006). It has also been proposed that the NC-derived visceral nervous system that innervates

the ureter may also modulate the frequency and amplitude of waves of contraction of the SMC layer, although very little is known about how pyeloureteral innervation is generated and how NC-derived neurons interact with the SMC contractile layer (Bohnenpoll and Kispert, 2014).

Consistent with the importance of the SMC layer and pacemaker cells in ureter peristalsis, defects in growth and differentiation of the ureteric mesenchyme (the precursor of the multilayered ureter) or defects in development of pacemaker cells can lead to congenital hydronephrosis (Bohnenpoll and Kispert, 2014; Cain et al., 2011). However, only a few regulators of ureteric mesenchyme growth and differentiation have been discovered so far, including *Six1*, *Tbx18* and *Sox9*, and with canonical Wnt, BMP and Shh pathways being the main signaling inputs implicated in the regulation of these genes (Bohnenpoll and Kispert, 2014). The tissue origin of pacemaker cells and the molecular mechanisms that control pacemaker cell-mediated peristalsis activity are even less well understood. So far, only Shh, angiotensin and calcineurin signaling have been implicated as regulators of these processes (Bohnenpoll and Kispert, 2014; Cain et al., 2011).

The severe hydronephrosis phenotype of *R26(P3Cre)*$^{EnRCdx1/+}$ mice (Figure 4.3) suggests that Cdx proteins may be important early regulators of ureter morphogenesis. Ureter develops from e11.5 onwards, from condensation and differentiation of the ureteric mesenchyme that surrounds the distal ureteric bud at the level of the cloaca (the primordium of bladder and urethra) in the hindlimb region (Bohnenpoll and Kispert, 2014). By performing DiI labeling tracing Herzlinger's lab demonstrated that the ureteric mesenchyme population derives from the tailbud, and that signals from the tailbud, particularly BMP could impact in ureter morphogenesis (Brenner-Anantharam et al., 2007). Given that Cdx transcripts mostly localize to the posterior growth zone of the tailbud at this stage (van de Ven et al., 2011), our main hypothesis is that Cdx proteins control the specification/growth of the pool of ureteric

mesenchymal progenitors that later localize to the cloaca region and thus before ureter formation *per se*. This hypothesis is supported by the observation that hydronephrosis has been observed with both the *P3Pro*-Cre (Figure2) and the *T*-Cre lines (data not shown). Indeed, both of these lines drive Cre-mediated expression of EnRCdx1 in the tailbud. Importantly, our hypothesis is also supported by studies proposing that Cdx impact in cloaca development and particularly in uro-rectal septum generation by acting earlier at the posterior growth zone (van de Ven et al., 2011). However, we also cannot rule out the possibility that Cdx proteins act via the NC to control development of ureteric pacemaker cells, since some unpublished data from N. Rosemblum's lab (personal communication of Dr. Maxime Bouchard) suggest that these cells are of NC origin.

Interestingly, unilateral hydronephrosis has been previously described in a subset of female mice over-expressing Cdx1 (OE1) (Gaunt and Paul, 2011), suggesting that Cdx dosage is important for ureter development. On the other hand, $Cdx2^{+/-}$::$Cdx4^{-/-}$ double mutant embryos exhibit uro-rectal septation defects that lead to recto-urinary fistula and mega-bladder but appear to not exhibit hydronephrosis (van de Ven et al., 2011). This suggests that hydronephrosis has been masked by Cdx functional overlap in these mutants and that the EnRCdx1 approach yield enough Cdx loss-of-function to unmask this phenotype. In $R26(P3Cre)^{EnRCdx1/+}$ mice, hydronephrosis is present in both sexes, and either unilaterally or bilaterally (Figure 4.3). We have detected this phenotype at e18.5, suggesting a developmental origin. We have also observed mega-bladder but at a very low penetrance.

Our preliminary characterization of the $R26(P3Cre)^{EnRCdx1/+}$ hydronephrosis phenotype suggests that there is a problem in ureter peristalsis. We have performed time-lapse live imaging in control and $R26(P3Cre)^{EnRCdx1/+}$::*Cdx1*KO mice at P5 (Figure 4.4), and detected problems with the proximo-distal coordination of ureter contractions in the mutants. The frequency of peristaltic waves appears to be lower in

the mutants and, more interestingly, we observed the formation of aberrant waves of contractions that oppose the normal proximal to distal wave of contractions (Figure 4.4). These aberrant waves most likely propel the urine back to the renal pelvis and thereby explain the urine accumulation in the kidney and hydronephrosis. These observations point to defects in the pacemaker machinery rather than obstruction. Further studies will be needed for validating these results in conventional genetic Cdx loss-of-function models, as well as for uncovering the molecular mechanisms involved. The main questions are: 1) Is this obstructive or non-obstructive hydronephrosis?; 2) Which is the tissue origin of the phenotype: NC or ureteric mesenchyme?; 3) What are the cell populations affected: SMC or pacemaker cells?; and 4) What is the problem with the affected cells?

4.7 Impact of the work

4.7.1 The conditional EnRCdx1 mouse model as a tool to unmask Cdx functions in the mouse

The EnRCdx1 mouse model generated in this work can be combined with a battery of Cre lines to direct temporal and tissue specific inactivation of Cdx functions, thereby circumventing functional overlap and/or early embryonic death. This model can be useful to unmask Cdx functions during embryonic development, adult tissues (intestine) or cancer. For example, it would be interesting to use this model to study the Cdx role in murine haematopoiesis. Indeed, *Cdx* genes have been implicated as important regulators of haematopoiesis in zebrafish (Davidson et al., 2003; Davidson and Zon, 2006) but similar studies in mice have been hampered by functional overlap, early lethality and severe truncation phenotypes of *Cdx* mutant models (Lengerke and Daley, 2012). A critical Cdx role is expected in murine hematopoiesis as studies with *in vitro*-differentiated ES cells have shown that Cdx proteins activate blood specification by relaying information from canonical Wnt and BMP signaling to key *Hox* promoters (Lengerke et al., 2007; Lengerke et al., 2008).

4.7.2 Cdx and neural development

As defined by Robert Bolande in 1974, neurocristopathies are a group of disorders including tumors, tumor syndromes and malformations caused by abnormal development of NCC (Bolande, 1997). Neurocristopathies frequently involve pigmentary anomalies and may be associated with NTDs, which represent the second most prevalent group of human congenital malformations after congenital heart defects (Copp et al., 2003; Mitchell, 2005; Nye et al., 1999). In spite of its clinical importance, only a few genes have been involved in the etiology of these diseases in human, including *PAX3*. A strict control of *Pax3* expression during embryogenesis (where, when and how long *Pax3* is expressed) is critical for normal development of the nervous system. Downregulation of neural *Pax3* expression in the mouse, as seen in Splotch mutants, leads to several NC and NT disorders that mimic some of the defects observed in humans, which include Waardenburg syndrome (WS) types I and III (that may also be associated with aganglionic megacolon), cardiac defects and *spina bifida* (Agopian et al., 2013; Mathieu et al., 1990; Tassabehji et al., 1992). Because of this clinical relevance, a full comprehension of the mechanisms that regulate *Pax3* neural expression and more generally, NC and NT development is mandatory. In the present work, we have contributed to this effort by showing that the small regulatory sequence of the *Pax3* locus NCE2 is sufficient to recapitulate *Pax3* expression in the mouse caudal neuroectoderm. We have also identified Cdx and Zic2 as two novel regulators of *Pax3* expression and shown that these transcription factors are at the core of a regulatory circuit that integrates signals from A-P and D-V instructive cues on NCE2. To extend these results, we have further shown that Cdx proteins directly control the expression of *Msx1* and *FoxD3* which are other important regulators of the trunk NC-GRN. Moreover, we have demonstrated that Cdx function is required for normal trunk NC development as single allelic expression of a Cdx dominant negative protein in mice led to pigmentary anomalies and a reduced number of myenteric neural ganglia. In addition, P3proCre-directed EnRCdx1 expression

affects NT development, consistent with previous work evidencing a critical Cdx role in the control of NT closure (Savory et al., 2011a). Taken together, our work in the mouse has identified Cdx proteins as central regulators of nervous system development, suggesting their involvement in the etiology of neurocristopathies and NTDs in humans. Their strategic position at the top of the trunk NC-GRN and their essential role in NT development make *Cdx* genes very interesting candidates to use in stem cell and gene therapies aimed at repairing spinal neural tissues affected by disease or injury.

4.7.3 Cdx from development to cancer?

This work is but another example that Cdx proteins are *bona fide* pioneer factors, which act to start patterning and posterior development in all germ layers (mesoderm, endoderm, neuoroectoderm and NC) and are then turned off once they have done their work. Reactivation and/or ectopic *Cdx* gene expression in adult tissues is associated with malignant transformation of cells (cancer) (Maulbecker and Gruss 1993; Soubeyran et al., 2001). Particularly, aberrant activation of *Cdx* has been associated with development of the precancerous intestinal metaplasia in the stomach and esophagus as well as in acute myeloid and lymphoid leukemias (reviewed in Guo et al., 2004; Lengerke and Daley, 2012) . In these malignant cells, it is believed that Cdx proteins reactivate the same pathways they normally control during development, for example, re-expression of *Hox* genes in leukemia (Lengerke and Daley, 2012). It is important to point out that Cdx1 and Cdx2 have also been found to act as tumor supressors in the small intestine and colon, in which they are normally expressed during adulthood (Grainger et al., 2013; Hryniuk et al., 2014).

Melanoma results from the malignant transformation of melanocytes. It is considered the most aggressive human skin cancer and one of the most metastatic and fatal types of all cancers (Bastian, 2014; Lang et al., 2013; Levy et al., 2006). Melanoma is a classical example of cancer caused by genetic predisposition and exposition to

environmental factors (solar UV radiation). Despite massive sun protection campaigns, both the number of cases and associated mortality is increasing globally (~48 000 deaths/year), and there is currently no cure for this lethal disease (Lang et al., 2013; Maguire et al., 2014). It is extensively recognised that the RAS/RAF/MEK/ERK1/2 pathway is hyperactivated in melanoma, with activating oncogenic mutations in the RAF member and serine/threonine kinase BRAF been found in 40-50% of cutaneous melanomas. Although the use of RAF and MEK inhibitors has been shown to suppress tumor growth and prolong survival of BRAF-mutant melanoma patients, approximately 50% of the patients develop resistance to the treatment after 6-7 months (Abel et al., 2013). A better understanding of the molecular mechanisms implicated in the malignant transformation of melanocytes and the identification of novel molecular effectors for use in targeted therapies is thus mandatory.

The work presented in this thesis has uncovered a new Cdx role in the control of melanocyte development. Since it is well recognized that the metastasis and escape from apoptosis of melanoma cells result from reactivation of the NC and melanocyte developmental program (Maguire et al., 2014), it is tempting to speculate that Cdx proteins may be implicated in the development of human melanoma as well. Canonical Wnt signaling, one the major pathways involved in *Cdx* activation during embryonic development, and the Cdx target genes revealed in this work (*Pax3*, *Foxd3* and *Sox9*) are all important players in melanoma (Medic and Ziman, 2009; Sinnberg et al., 2011). Hence, CDX proteins may also act downstream of canonical WNT signaling to start the melanoma developmental program. *PAX3* expression is reactivated in melanoma, where it acts not only as a *bona fide* marker but also as a survival factor to inhibit apoptosis via the same mechanism that are used during embryonic development (He et al., 2005; Medic and Ziman, 2009). Consistent with these roles, high levels of *PAX3* expression are indicative of more aggressive melanoma, whereas reduction in expression is associated with a good prognostic and

tumor apoptosis (reviewed in Kubic et al., 2008). On the other hand, *FOXD3* expression is also upregulated in melanoma, in which this stemness factor has been shown to either inhibit or activate mutant BRAF-melanoma progression. More specifically, it has been found that *FOXD3* expression is induced upon BRAF/MEK pathway inhibition and that FOXD3 promotes melanoma cell viability and resistance to treatment, by directly activating the expression of *ERRB3* and consequently promoting ERRB3 and AKT signaling (Abel et al., 2013). By contrast, other work has revealed that *FOXD3* upregulation abrogates the metastatic potential of BRAF-melanoma, in which case FOXD3 was shown to directly repress the expression of the pro-metastatic gene *TWIST1* (Weiss et al., 2014a). Upregulation of *SOX9* in human and mouse melanoma cell lines induces melanoma cell arrest and restores sensitivity to RA treatment (Passeron et al., 2009). Thus, by indirectly downregulating *Sox9* expression levels, Cdx proteins may also facilitate melanoma progression and resistance to treatment. Importantly, gene expression data mining with the UCSC Cancer Genomics Browser (https://genome-cancer.ucsc.edu/proj/site/hgHeatmap/) suggests that *Cdx1*, *Cdx2* and/or *Cdx4* are expressed in approximately 35% of short-term melanoma culture samples, including BRAF-mutated melanomas (Lin et al., 2008). Taking into consideration all these evidences, it would be interesting to test whether the here proposed Cdx-dependent melanocyte/melanoma gene regulatory network is at work and how Cdx loss-of-function may impact on the progress of melanoma. The strategic position of *Cdx* genes upstream of *Pax3*, *FoxD3* and *Sox9* makes them very interesting candidates to concomitantly target several important melanoma regulators.

4.8 Figures

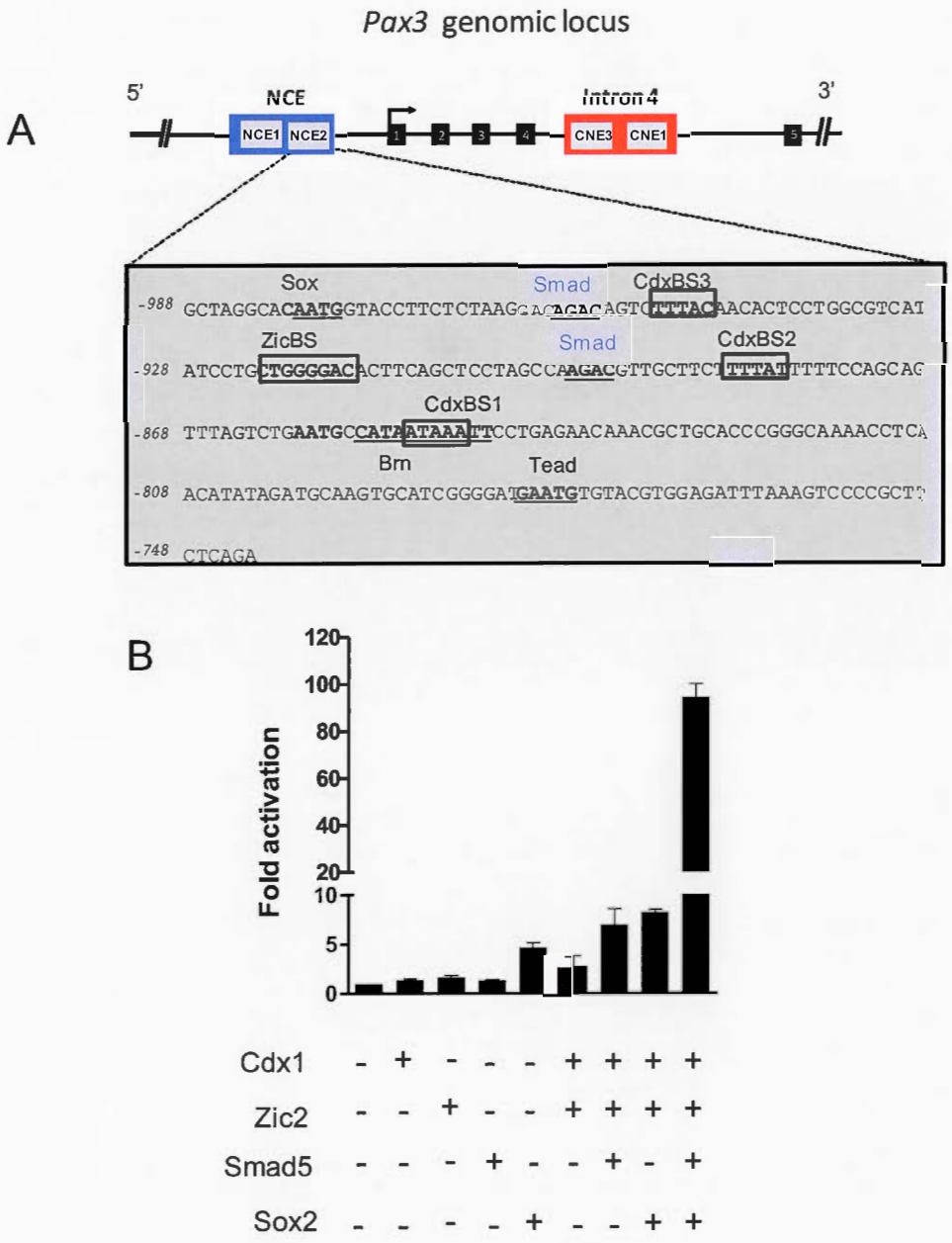


Figure 4.1 Canonical BMP-Smad signaling is a candidate as being involved in the restriction of the *Pax3* expression domain to the posterior neural plate border by interacting with the Cdx-Zic2-Sox2 complex via NCE2.

(A) Schematic representation of the *Pax3* genomic locus showing the relative position of *Pax3* regulatory regions and snapshot of the NCE2 sequence as published in (Sanchez-Ferras et al., 2014). The two putative Smad binding sites according to MatInspector binding site analysis are denoted in blue are located in close proximity to the three functional Cdx, the Zic binding site (in boxes), and the putative Sox binding site (underlined). (B) Co-transfection analysis in N2a cells. Cells were transiently transfected with the *Pax3*NCE2-luc reporter alone or co-transfected with $_{FLAG}$Cdx1 (5ng), $_{HA}$Zic2 (10 ng), Smad5 (75 ng) and MycSox2 (50 ng) expression vector alone or in combination, total amount of DNA 150 ng per well. Note that these experimental conditions differ from those used in (Sanchez-Ferras et al., 2014) when evaluating the synergistic interaction between Cdx1, Zic2 and Sox2. Results are presented as fold activation compared to the relevant reporter vector alone. n=4-14 independent experiments performed in triplicate. Error bars denote s.e.m. Note the strong synergistic activation of *Pax3*NCE2 when Smad5 is co-expressed with Cdx1, Zic2 and Sox2.

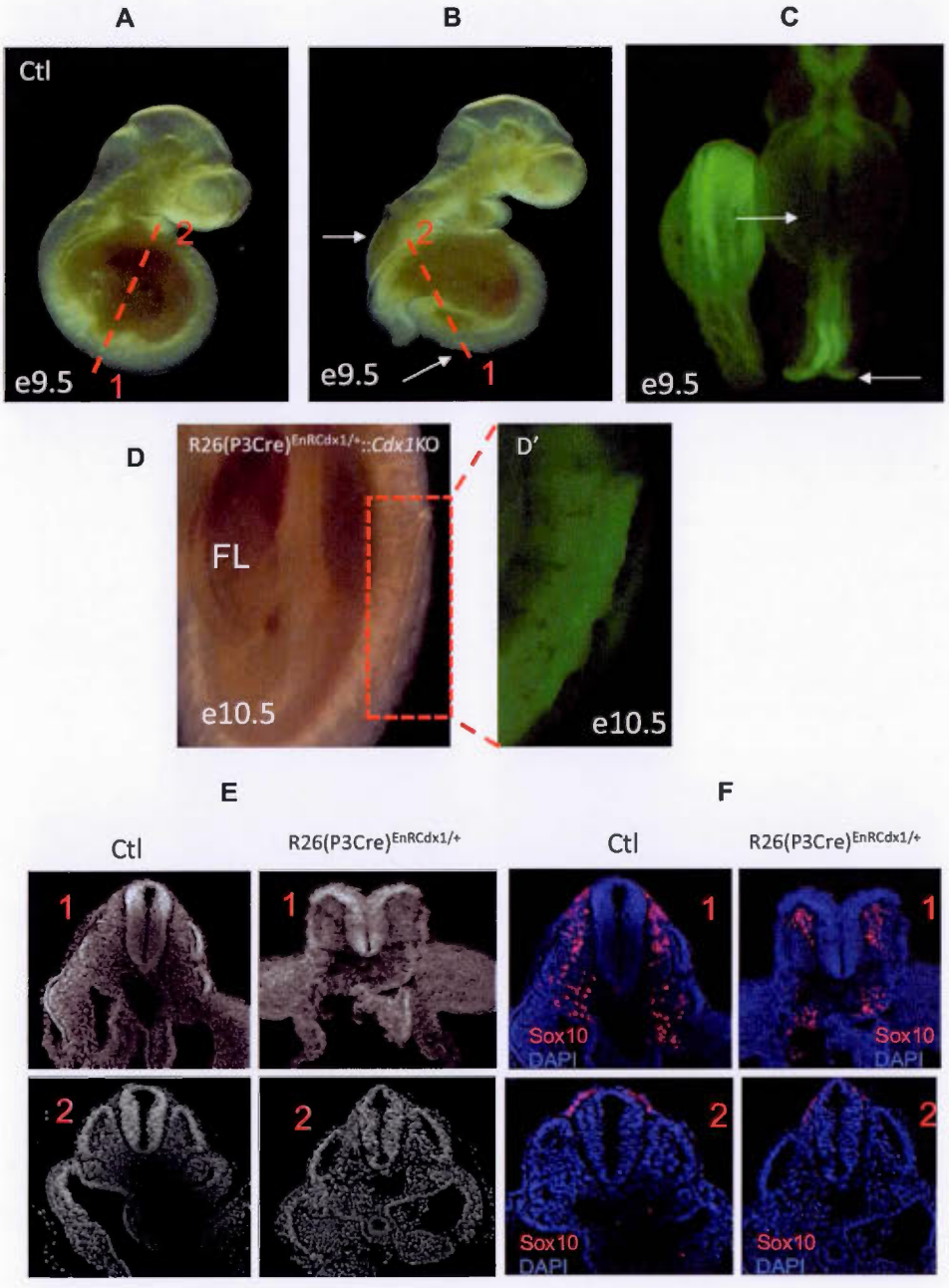


Figure 4.2 R26(P3Cre)$^{EnRCdx1/+}$ e9.5 embryos exhibit NT malformations.

(A,B) Bright field lateral image of Ctl (R26R$^{EnRCdx1/+}$,A) and R26(P3Cre)$^{EnRCdx1/+}$ (B) embryos. (C) Dorsal view and fluorescence (EGFP) image of theR26(P3Cre)$^{EnRCdx1/+}$ embryo

shown in B. Note the presence of open neural tube defects at the level of the hindbrain-spinal cord boundary and forelimb in the mutant (white arrows in B and C). The red dashed lines and numbers represent the level at which the sections shown in E and F were taken. (D) Bright field view of the forelimb (FL) region of a R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO e10.5 embryo with an open NT lesion (denoted in a dashed red box). (D') High magnification view of the lesion showing EGFP fluorescence. (E) Vibratome sections at the forelimb level (1) and more posterior levels (2) in Ctl and R26(P3Cre)$^{EnRCdx1/+}$ embryos shown in A and B. Note that the dorsal NT fails to close in the mutant at the forelimb level (1) whereas at more posterior regions (2) the NT is closed but irregularly shaped in the mutant. (F) Sox10 immunofluorescence analysis of the NC population in these regions shows that the presence of NT malformations in R26(P3Cre)$^{EnRCdx1/+}$ do not affect delamination of NC progenitors.

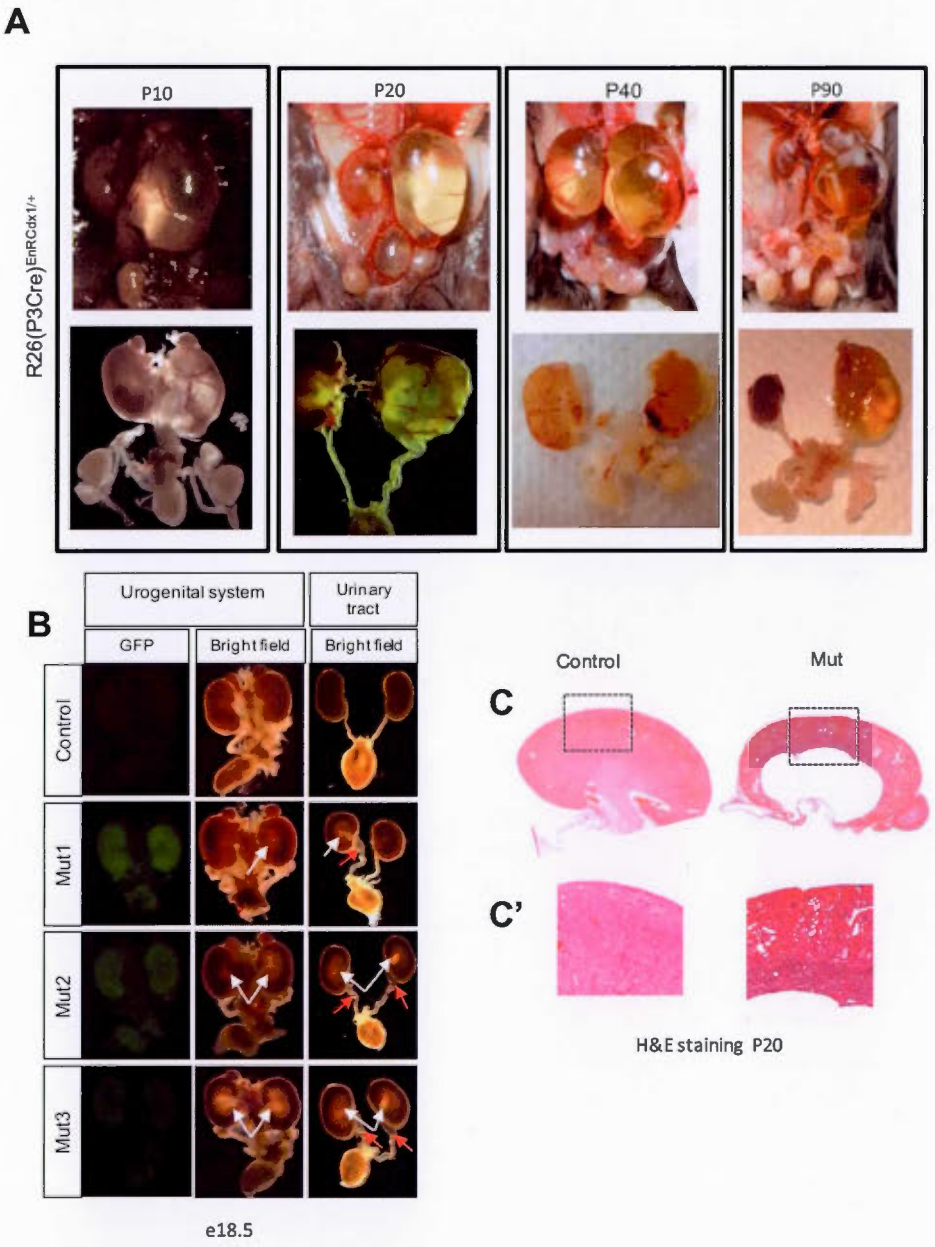


Figure 4.3 R26(P3Cre)$^{EnRCdx1/+}$ double transgenic mice exhibit hydronephrosis as early as e18.5.

Analysis of the urinary system of R26(P3Cre)$^{EnRCdx1/+}$ mice at different stages: (A) at postnatal day (P) 10, P20, P40, P90 and (B) peri-natal (e18.5). MUT in B refers to a

R26(P3Cre)$^{EnRCdx1/+}$ genotype. Note that mutant mice exhibit hydronephrosis (white arrows in B) and hydroureter (red arrows in B) with the kidney and ureter swollen in size. Also note that hydronephrosis may present both unilaterally affecting the left kidney or bilaterally. (C) H&E-stained paraffin sections from a P20 control and R26(P3Cre)$^{EnRCdx1/+}$ mice. Note the severe atrophy of the renal parenchyma, massive nephron loss and dilatation of the renal pelvis in the hydronephrotic kidney of the mutant.

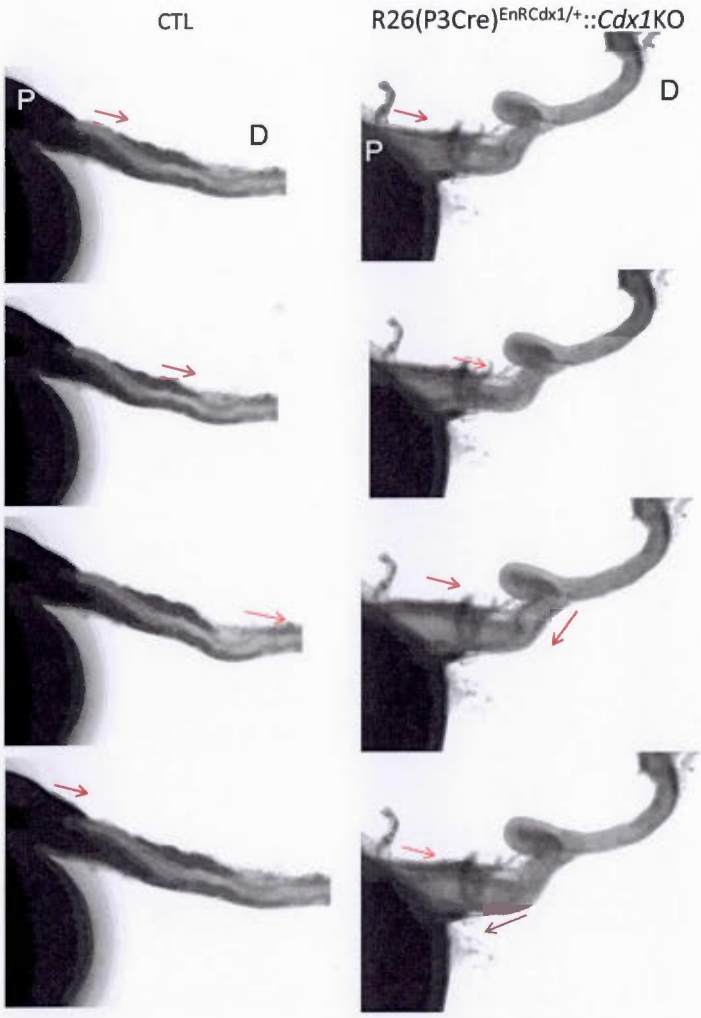


Figure 4.4 Defective ureter peristalsis in R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO mutants.

Serial images taken at 500 ms intervals during a peristaltic cycle in a control (*R26R*$^{EnRCdx1/+}$:: *Cdx1KO*) and a mutant R26(P3Cre)$^{EnRCdx1/+}$::*Cdx1*KO at postnatal day (P)5. Red arrows indicate the direction and location of the peristaltic wave. See the proximal (P)-to distal (D) wave of peristalsis in the control. In the mutant an additional aberrant peristaltic wave coming in a D to P sense propels the urine back to the renal pelvis.

CONCLUSIONS

This work showed the first *in vivo* evidence for a direct Cdx role in trunk NC development in the mouse. We have generated a new conditional (Cre-loxP dependent) pan-Cdx loss-of-function mouse model called (R26R-$_{FLAG}$EnRCdx1) that can be used to circumvent Cdx functional overlap and/or early embryonic lethality with the aim of uncovering novel Cdx functions in the mouse. Via targeted expression of the obligatory repressor fusion protein EnRCdx1 in the dorsal NT and NC, we showed that Cdx function is required for normal NC-derived melanocyte and myenteric neural ganglia development, with Cdx loss-of-function leading to aberrant melanoblast development at embryonic stage, hypo-pigmentation in postnatal life, as well as reduced number of myenteric neural ganglia in the distal colon of adult mice. We have also observed defects in NT development, consistent with previous work reporting a critical Cdx role in the control of NT closure and morphogenesis (Savory et al., 2011a; van de Ven et al., 2011). To extend the knowledge of Cdx functions in the mouse, we also showed that Cdx proteins are involved in the development of the urogenital system *in vivo*, as revealed by the severe hydronephrosis observed upon Cdx knockdown targeted to either neural (*P3Pro*-Cre line) or mesoderm (*T*-Cre line) tissues. This led us suggesting that a defective Cdx function may be implicated in the etiology of hydronephrosis, and proposing that the EnRCdx1 mouse line is a good model to better understand the molecular mechanisms implicated in this frequent and serious disease.

We demonstrated that *Cdx* genes occupy a strategic position at the top of the trunk NC-GRN downstream of the signaling inputs and upstream of the important regulators *Pax3*, *Msx1* and *Foxd3*. Mechanistic studies revealed that Cdx exhibit a common *modus operandi* and directly regulate expression of *Pax3*, *Msx1* and *Foxd3* via binding to conserved elements within the proximal promoter region. Particularly, we have extensively characterized the mechanism by which Cdx proteins regulate

Pax3 neural expression and narrowed down the minimal Cdx responsive regulatory region to a short (~250 bp in length) trunk NC enhancer located within the 1.6 kb *Pax3* proximal promoter and called NCE2. Our work showed for the first time that *Pax3* NCE2 behaves as a trunk NC "super-enhancer" that recapitulate induction and dorsal restriction of *Pax3* expression in the mouse caudal neuroectoderm. We demonstrated that Cdx directly regulate *Pax3* expression via binding to CdxBS present in NCE2. By performing perturbation analysis of Cdx function and canonical Wnt signaling in N2a and P19 cell lines, we demonstrated that Cdx proteins act as intermediates in the mechanism of Wnt-mediated induction of *Pax3* neural expression. Moreover, we have uncovered the regulatory circuit acting via NCE2. We identified the zinc finger domain Zic2 transcription factor as a direct *Pax3* NCE2 regulator and Cdx interacting partner and proposed a model in which the inductive A-P positional cue from the Wnt-Cdx pathway are integrated with a D-V instructive cue from the Shh-Nkx6.1-Zic2 pathway and a general neural input from Sox2 to ensure induction and dorsal restriction of *Pax3* expression in the caudal neuroectoderm.

Taken together with the work of others, this work demonstrated that Cdx are essential regulators of nervous system development and putative candidates as being involved in the etiology of neurocristopathies and NTDs in humans. Moreover, given the several evidences suggesting that cancer co-opts the mechanisms of NC development (Maguire et al., 2014), this work opens the possibility that Cdx are also implicated in development of tumors derived from the NC such as melanoma.

PERSPECTIVES

Future projects derived from this work:

1) Characterize the mechanism of regulation of *Pax3*NCE2 by BMP signaling and the integration of the BMP signaling input with the Cdx-dependent NCE2 regulatory circuit.

2) Study the regulation of *Pax3* expression by RA-Cdx and FGF-Cdx pathways as well as whether and how these inputs are integrated with the NCE2 regulatory circuit.

3) Study the mechanisms by which Cdx proteins regulate *Msx1* and *FoxD3* NC expression.

4) Study the role of Cdx proteins in urogenital system development and their implication in congenital anomalies of the kidney and urinary tract (CAKUT).

5) Obtain a more complete understanding of the Cdx role in trunk NC-development via characterization of the EnRCdx1 homozygous phenotype.

6) Analyze Cdx functions in melanocyte / melanoma development.

REFERENCES

Abel, E. V., Basile, K. J., Kugel, C. H., 3rd, Witkiewicz, A. K., Le, K., Amaravadi, R. K., Karakousis, G. C., Xu, X., Xu, W., Schuchter, L. M. et al. (2013). Melanoma adapts to RAF/MEK inhibitors through FOXD3-mediated upregulation of ERBB3. *J Clin Invest* 123, 2155-68.

Abitua, P. B., Wagner, E., Navarrete, I. A. and Levine, M. (2012). Identification of a rudimentary neural crest in a non-vertebrate chordate. *Nature* 492, 104-7.

Abzhanov, A., Tzahor, E., Lassar, A. B. and Tabin, C. J. (2003). Dissimilar regulation of cell differentiation in mesencephalic (cranial) and sacral (trunk) neural crest cells in vitro. *Development* 130, 4567-79.

Adameyko, I., Lallemend, F., Aquino, J. B., Pereira, J. A., Topilko, P., Muller, T., Fritz, N., Beljajeva, A., Mochii, M., Liste, I. et al. (2009). Schwann cell precursors from nerve innervation are a cellular origin of melanocytes in skin. *Cell* 139, 366-79.

Agopian, A. J., Bhalla, A. D., Boerwinkle, E., Finnell, R. H., Grove, M. L., Hixson, J. E., Shimmin, L. C., Sewda, A., Stuart, C., Zhong, Y. et al. (2013). Exon sequencing of PAX3 and T (brachyury) in cases with spina bifida. *Birth Defects Res A Clin Mol Teratol* 97, 597-601.

Alappat, S., Zhang, Z. Y. and Chen, Y. P. (2003). Msx homeobox gene family and craniofacial development. *Cell Res* 13, 429-42.

Allan, D., Houle, M., Bouchard, N., Meyer, B. I., Gruss, P. and Lohnes, D. (2001). RARgamma and Cdx1 interactions in vertebral patterning. *Dev Biol* 240, 46-60.

Allen, B. L., Tenzen, T. and McMahon, A. P. (2007). The Hedgehog-binding proteins Gas1 and Cdo cooperate to positively regulate Shh signaling during mouse development. *Genes Dev* 21, 1244-57.

Alvarez Martinez, C. E., Binato, R., Gonzalez, S., Pereira, M., Robert, B. and Abdelhay, E. (2002). Characterization of a Smad motif similar to Drosophila mad in the mouse Msx 1 promoter. *Biochem Biophys Res Commun* 291, 655-62.

Amiel, J., Sproat-Emison, E., Garcia-Barcelo, M., Lantieri, F., Burzynski, G., Borrego, S., Pelet, A., Arnold, S., Miao, X., Griseri, P. et al. (2008). Hirschsprung disease, associated syndromes and genetics: a review. *J Med Genet* 45, 1-14.

Ang, S. L., Conlon, R. A., Jin, O. and Rossant, J. (1994). Positive and negative signals from mesoderm regulate the expression of mouse Otx2 in ectoderm explants. *Development* 120, 2979-89.

Araya, C., Tawk, M., Girdler, G. C., Costa, M., Carmona-Fontaine, C. and Clarke, J. D. (2014). Mesoderm is required for coordinated cell movements within zebrafish neural plate in vivo. *Neural Dev* 9, 9.

Arduini, B. L. and Brivanlou, A. H. (2012). Modulation of FOXD3 activity in human embryonic stem cells directs pluripotency and paraxial mesoderm fates. *Stem Cells* 30, 2188-98.

Arkell, R. M. and Tam, P. P. (2012). Initiating head development in mouse embryos: integrating signalling and transcriptional activity. *Open Biol* 2, 120030.

Arnold, S. J. and Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Biol* 10, 91-103.

Aruga, J. (2004). The role of Zic genes in neural development. *Mol Cell Neurosci* 26, 205-21.

Aruga, J., Inoue, T., Hoshino, J. and Mikoshiba, K. (2002a). Zic2 controls cerebellar development in cooperation with Zic1. *J Neurosci* 22, 218-25.

Aruga, J., Kamiya, A., Takahashi, H., Fujimi, T. J., Shimizu, Y., Ohkawa, K., Yazawa, S., Umesono, Y., Noguchi, H., Shimizu, T. et al. (2006). A wide-range phylogenetic analysis of Zic proteins: implications for correlations between protein structure conservation and body plan complexity. *Genomics* 87, 783-92.

Aruga, J., Minowa, O., Yaginuma, H., Kuno, J., Nagai, T., Noda, T. and Mikoshiba, K. (1998). Mouse Zic1 is involved in cerebellar development. *J Neurosci* 18, 284-93.

Aruga, J., Mizugishi, K., Koseki, H., Imai, K., Balling, R., Noda, T. and Mikoshiba, K. (1999). Zic1 regulates the patterning of vertebral arches in cooperation with Gli3. *Mech Dev* 89, 141-50.

Aruga, J., Nagai, T., Tokuyama, T., Hayashizaki, Y., Okazaki, Y., Chapman, V. M. and Mikoshiba, K. (1996). The mouse zic gene family. Homologues of the Drosophila pair-rule gene odd-paired. *J Biol Chem* 271, 1043-7.

Aruga, J., Tohmonda, T., Homma, S. and Mikoshiba, K. (2002b). Zic1 promotes the expansion of dorsal neural progenitors in spinal cord by inhibiting neuronal differentiation. *Dev Biol* 244, 329-41.

Aruga, J., Yokota, N., Hashimoto, M., Furuichi, T., Fukuda, M. and Mikoshiba, K. (1994). A novel zinc finger protein, zic, is involved in neurogenesis, especially in the cell lineage of cerebellar granule cells. *J Neurochem* 63, 1880-90.

Auerbach, R. (1954). Analysis of the developmental effects of a lethal mutation in the house mouse. *J Exp Zoology* 127, 305–329.

Avsian-Kretchmer, O. and Hsueh, A. J. (2004). Comparative genomic analysis of the eight-membered ring cystine knot-containing bone morphogenetic protein antagonists. *Mol Endocrinol* 18, 1-12.

Ayme, S. and Philip, N. (1995). Possible homozygous Waardenburg syndrome in a fetus with exencephaly. *Am J Med Genet* 59, 263-5.

Baba, Y., Nosho, K., Shima, K., Freed, E., Irahara, N., Philips, J., Meyerhardt, J. A., Hornick, J. L., Shivdasani, R. A., Fuchs, C. S. et al. (2009). Relationship of CDX2 loss with molecular features and prognosis in colorectal cancer. *Clin Cancer Res* 15, 4665-73.

Bachiller, D., Klingensmith, J., Kemp, C., Belo, J. A., Anderson, R. M., May, S. R., McMahon, J. A., McMahon, A. P., Harland, R. M., Rossant, J. et al. (2000). The organizer factors Chordin and Noggin are required for mouse forebrain development. *Nature* 403, 658-61.

Balaskas, N., Ribeiro, A., Panovska, J., Dessaud, E., Sasai, N., Page, K. M., Briscoe, J. and Ribes, V. (2012). Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. *Cell* 148, 273-84.

Baldwin, C. T., Hoth, C. F., Amos, J. A., da-Silva, E. O. and Milunsky, A. (1992). An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. *Nature* 355, 637-8.

Baldwin, C. T., Hoth, C. F., Macina, R. A. and Milunsky, A. (1995). Mutations in PAX3 that cause Waardenburg syndrome type I: ten new mutations and review of the literature. *Am J Med Genet* 58, 115-22.

Bang, A. G., Papalopulu, N., Goulding, M. D. and Kintner, C. (1999). Expression of Pax-3 in the lateral neural plate is dependent on a Wnt-mediated signal from posterior nonaxial mesoderm. *Dev Biol* 212, 366-80.

Bang, A. G., Papalopulu, N., Kintner, C. and Goulding, M. D. (1997). Expression of Pax-3 is initiated in the early neural plate by posteriorizing signals produced by the organizer and by posterior non-axial mesoderm. *Development* 124, 2075-85.

Bani-Yaghoub, M., Tremblay, R. G., Lei, J. X., Zhang, D., Zurakowski, B., Sandhu, J. K., Smith, B., Ribecco-Lutkiewicz, M., Kennedy, J., Walker, P. R. et al. (2006). Role of Sox2 in the development of the mouse neocortex. *Dev Biol* 295, 52-66.

Barber, T. D., Barber, M. C., Cloutier, T. E. and Friedman, T. B. (1999). PAX3 gene structure, alternative splicing and evolution. *Gene* 237, 311-9.

Barembaum, M. and Bronner-Fraser, M. (2005). Early steps in neural crest specification. *Semin Cell Dev Biol* 16, 642-6.

Barros, R., da Costa, L. T., Pinto-de-Sousa, J., Duluc, I., Freund, J. N., David, L. and Almeida, R. (2011). CDX2 autoregulation in human intestinal metaplasia of the stomach: impact on the stability of the phenotype. *Gut* 60, 290-8.

Barros, R., Pereira, B., Duluc, I., Azevedo, M., Mendes, N., Camilo, V., Jacobs, R. J., Paulo, P., Santos-Silva, F., van Seuningen, I. et al. (2008). Key elements of the BMP/SMAD pathway co-localize with CDX2 in intestinal metaplasia and regulate CDX2 expression in human gastric cell lines. *J Pathol* 215, 411-20.

Bastian, B. C. (2014). The molecular pathology of melanoma: an integrated taxonomy of melanocytic neoplasia. *Annu Rev Pathol* 9, 239-71.

Baxter, L. L. and Pavan, W. J. (2013). The etiology and molecular genetics of human pigmentation disorders. *Wiley Interdiscip Rev Dev Biol* 2, 379-92.

Baynash, A. G., Hosoda, K., Giaid, A., Richardson, J. A., Emoto, N., Hammer, R. E. and Yanagisawa, M. (1994). Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons. *Cell* 79, 1277-85.

Beachy, P. A., Hymowitz, S. G., Lazarus, R. A., Leahy, D. J. and Siebold, C. (2010). Interactions between Hedgehog proteins and their binding partners come into view. *Genes Dev* 24, 2001-12.

Beck, F. (2002). Homeobox genes in gut development. *Gut* 51, 450-4.

Beck, F., Erler, T., Russell, A. and James, R. (1995). Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Dev Dyn* 204, 219-27.

Beddington, R. S. (1994). Induction of a second neural axis by the mouse node. *Development* 120, 613-20.

Beddington, R. S. and Robertson, E. J. (1999). Axis development and early asymmetry in mammals. *Cell* 96, 195-209.

Bel-Vialar, S., Itasaki, N. and Krumlauf, R. (2002). Initiating Hox gene expression: in the early chick neural tube differential sensitivity to FGF and RA signaling subdivides the HoxB genes in two distinct groups. *Development* 129, 5103-15.

Beland, M. and Lohnes, D. (2005). Chicken ovalbumin upstream promoter-transcription factor members repress retinoic acid-induced Cdx1 expression. *J Biol Chem* 280, 13858-62.

Beland, M., Pilon, N., Houle, M., Oh, K., Sylvestre, J. R., Prinos, P. and Lohnes, D. (2004). Cdx1 autoregulation is governed by a novel Cdx1-LEF1 transcription complex. *Mol Cell Biol* 24, 5028-38.

Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., Naef, F., Robertson, E. J. and Constam, D. B. (2006). The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. *Dev Cell* 11, 313-23.

Benahmed, F., Gross, I., Gaunt, S. J., Beck, F., Jehan, F., Domon-Dell, C., Martin, E., Kedinger, M., Freund, J. N. and Duluc, I. (2008). Multiple regulatory regions control the complex expression pattern of the mouse Cdx2 homeobox gene. *Gastroenterology* 135, 1238-1247, 1247 e1-3.

Bergsland, M., Ramskold, D., Zaouter, C., Klum, S., Sandberg, R. and Muhr, J. (2011). Sequentially acting Sox transcription factors in neural lineage development. *Genes Dev* 25, 2453-64.

Betancur, P., Bronner-Fraser, M. and Sauka-Spengler, T. (2010). Assembling neural crest regulatory circuits into a gene regulatory network. *Annu Rev Cell Dev Biol* 26, 581-603.

Betters, E., Liu, Y., Kjaeldgaard, A., Sundstrom, E. and Garcia-Castro, M. I. (2010). Analysis of early human neural crest development. *Dev Biol* 344, 578-92.

Blank, M. C., Grinberg, I., Aryee, E., Laliberte, C., Chizhikov, V. V., Henkelman, R. M. and Millen, K. J. (2011). Multiple developmental programs are altered by loss of Zic1 and Zic4 to cause Dandy-Walker malformation cerebellar pathogenesis. *Development* 138, 1207-16.

Blumberg, B., Wright, C. V., De Robertis, E. M. and Cho, K. W. (1991). Organizer-specific homeobox genes in Xenopus laevis embryos. *Science* 253, 194-6.

Bohnenpoll, T. and Kispert, A. (2014). Ureter growth and differentiation. *Semin Cell Dev Biol.*

Bolande, R. P. (1997). Neurocristopathy: its growth and development in 20 years. *Pediatr Pathol Lab Med* 17, 1-25.

Bondurand, N., Pingault, V., Goerich, D. E., Lemort, N., Sock, E., Le Caignec, C., Wegner, M. and Goossens, M. (2000). Interaction among SOX10, PAX3 and MITF, three genes altered in Waardenburg syndrome. *Hum Mol Genet* 9, 1907-17.

Bonner, C. A., Loftus, S. K. and Wasmuth, J. J. (1995). Isolation, characterization, and precise physical localization of human CDX1, a caudal-type homeobox gene. *Genomics* 28, 206-11.

Borycki, A. G., Li, J., Jin, F., Emerson, C. P. and Epstein, J. A. (1999). Pax3 functions in cell survival and in pax7 regulation. *Development* 126, 1665-74.

Bouchard, M., Grote, D., Craven, S. E., Sun, Q., Steinlein, P. and Busslinger, M. (2005). Identification of Pax2-regulated genes by expression profiling of the mid-hindbrain organizer region. *Development* 132, 2633-43.

Bovolenta, P., Esteve, P., Ruiz, J. M., Cisneros, E. and Lopez-Rios, J. (2008). Beyond Wnt inhibition: new functions of secreted Frizzled-related proteins in development and disease. *J Cell Sci* 121, 737-46.

Brennan, J., Lu, C. C., Norris, D. P., Rodriguez, T. A., Beddington, R. S. and Robertson, E. J. (2001). Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* 411, 965-9.

Brenner-Anantharam, A., Cebrian, C., Guillaume, R., Hurtado, R., Sun, T. T. and Herzlinger, D. (2007). Tailbud-derived mesenchyme promotes urinary tract segmentation via BMP4 signaling. *Development* 134, 1967-75.

Brewer, S., Feng, W., Huang, J., Sullivan, S. and Williams, T. (2004). Wnt1-Cre-mediated deletion of AP-2alpha causes multiple neural crest-related defects. *Dev Biol* 267, 135-52.

Briscoe, J. and Therond, P. P. (2013). The mechanisms of Hedgehog signalling and its roles in development and disease. *Nat Rev Mol Cell Biol* 14, 416-29.

Bronner-Fraser, M. and Stern, C. (1991). Effects of mesodermal tissues on avian neural crest cell migration. *Dev Biol* 143, 213-7.

Bronner, M. E. (2014). Migrating into Genomics with the Neural Crest. *Advances in Biology* 2014, 1-8.

Bronner, M. E. and LeDouarin, N. M. (2012). Development and evolution of the neural crest: an overview. *Dev Biol* 366, 2-9.

Brook, F. A. and Gardner, R. L. (1997). The origin and efficient derivation of embryonic stem cells in the mouse. *Proc Natl Acad Sci U S A* 94, 5709-12.

Brooke, N. M., Garcia-Fernandez, J. and Holland, P. W. (1998). The ParaHox gene cluster is an evolutionary sister of the Hox gene cluster. *Nature* 392, 920-2.

Brown, L. Y., Kottmann, A. H. and Brown, S. (2003). Immunolocalization of Zic2 expression in the developing mouse forebrain. *Gene Expr Patterns* 3, 361-7.

Brown, L. Y., Odent, S., David, V., Blayau, M., Dubourg, C., Apacik, C., Delgado, M. A., Hall, B. D., Reynolds, J. F., Sommer, A. et al. (2001). Holoprosencephaly due to mutations in ZIC2: alanine tract expansion mutations may be caused by parental somatic recombination. *Hum Mol Genet* 10, 791-6.

Brown, S. A., Warburton, D., Brown, L. Y., Yu, C. Y., Roeder, E. R., Stengel-Rutkowski, S., Hennekam, R. C. and Muenke, M. (1998). Holoprosencephaly due to mutations in ZIC2, a homologue of Drosophila odd-paired. *Nat Genet* 20, 180-3.

Bushdid, P. B., Chen, C. L., Brantley, D. M., Yull, F., Raghow, R., Kerr, L. D. and Barnett, J. V. (2001). NF-kappaB mediates FGF signal regulation of msx-1 expression. *Dev Biol* 237, 107-15.

Byrd, N. A. and Meyers, E. N. (2005). Loss of Gbx2 results in neural crest cell patterning and pharyngeal arch artery defects in the mouse embryo. *Dev Biol* 284, 233-45.

Cai, C. and Grabel, L. (2007). Directing the differentiation of embryonic stem cells to neural stem cells. *Dev Dyn* 236, 3255-66.

Cain, J. E., Islam, E., Haxho, F., Blake, J. and Rosenblum, N. D. (2011). GLI3 repressor controls functional development of the mouse ureter. *J Clin Invest* 121, 1199-206.

Cambray, N. and Wilson, V. (2002). Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* 129, 4855-66.

Cambray, N. and Wilson, V. (2007). Two distinct sources for a population of maturing axial progenitors. *Development* 134, 2829-40.

Carreira, A. C., Alves, G. G., Zambuzzi, W. F., Sogayar, M. C. and Granjeiro, J. M. (2014). Bone Morphogenetic Proteins: Structure, biological function and therapeutic applications. *Arch Biochem Biophys* 561C, 64-73.

Carrel, T., Purandare, S. M., Harrison, W., Elder, F., Fox, T., Casey, B. and Herman, G. E. (2000). The X-linked mouse mutation Bent tail is associated with a deletion of the Zic3 locus. *Hum Mol Genet* 9, 1937-42.

Catron, K. M., Wang, H., Hu, G., Shen, M. M. and Abate-Shen, C. (1996). Comparison of MSX-1 and MSX-2 suggests a molecular basis for functional redundancy. *Mech Dev* 55, 185-99.

Cavallo, R. A., Cox, R. T., Moline, M. M., Roose, J., Polevoy, G. A., Clevers, H., Peifer, M. and Bejsovec, A. (1998). Drosophila Tcf and Groucho interact to repress Wingless signalling activity. *Nature* 395, 604-8.

Chan, W. Y., Cheung, C. S., Yung, K. M. and Copp, A. J. (2004). Cardiac neural crest of the mouse embryo: axial level of origin, migratory pathway and cell autonomy of the splotch (Sp2H) mutant effect. *Development* 131, 3367-79.

Chang, C. P., McDill, B. W., Neilson, J. R., Joist, H. E., Epstein, J. A., Crabtree, G. R. and Chen, F. (2004). Calcineurin is required in urinary tract mesenchyme for the development of the pyeloureteral peristaltic machinery. *J Clin Invest* 113, 1051-8.

Chang, C. P., Shen, W. F., Rozenfeld, S., Lawrence, H. J., Largman, C. and Cleary, M. L. (1995). Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev* 9, 663-74.

Chang, C. P., Stankunas, K., Shang, C., Kao, S. C., Twu, K. Y. and Cleary, M. L. (2008). Pbx1 functions in distinct regulatory networks to pattern the great arteries and cardiac outflow tract. *Development* 135, 3577-86.

Charite, J., de Graaff, W., Consten, D., Reijnen, M. J., Korving, J. and Deschamps, J. (1998). Transducing positional information to the Hox genes: critical interaction of cdx gene products with position-sensitive regulatory elements. *Development* 125, 4349-58.

Charite, J., de Graaff, W., Shen, S. and Deschamps, J. (1994). Ectopic expression of Hoxb-8 causes duplication of the ZPA in the forelimb and homeotic transformation of axial structures. *Cell* 78, 589-601.

Chawengsaksophak, K. and Beck, F. (1996). Chromosomal localization of cdx2, a murine homologue of the Drosophila gene caudal, to mouse chromosome 5. *Genomics* 34, 270-1.

Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J. and Beck, F. (2004). Cdx2 is essential for axial elongation in mouse development. *Proc Natl Acad Sci U S A* 101, 7641-5.

Chawengsaksophak, K., James, R., Hammond, V. E., Kontgen, F. and Beck, F. (1997). Homeosis and intestinal tumours in Cdx2 mutant mice. *Nature* 386, 84-7.

Chen, J. K., Taipale, J., Cooper, M. K. and Beachy, P. A. (2002). Inhibition of Hedgehog signaling by direct binding of cyclopamine to Smoothened. *Genes Dev* 16, 2743-8.

Chen, W., ten Berge, D., Brown, J., Ahn, S., Hu, L. A., Miller, W. E., Caron, M. G., Barak, L. S., Nusse, R. and Lefkowitz, R. J. (2003). Dishevelled 2 recruits beta-arrestin 2 to mediate Wnt5A-stimulated endocytosis of Frizzled 4. *Science* 301, 1391-4.

Chen, Z. F. and Behringer, R. R. (1995). twist is required in head mesenchyme for cranial neural tube morphogenesis. *Genes Dev* 9, 686-99.

Cheung, M. and Briscoe, J. (2003). Neural crest development is regulated by the transcription factor Sox9. *Development* 130, 5681-93.

Cheung, M., Chaboissier, M. C., Mynett, A., Hirst, E., Schedl, A. and Briscoe, J. (2005). The transcriptional control of trunk neural crest induction, survival, and delamination. *Dev Cell* 8, 179-92.

Chevalier, R. L. (1998). Pathophysiology of obstructive nephropathy in the newborn. *Semin Nephrol* 18, 585-93.

Chiang, C., Litingtung, Y., Lee, E., Young, K. E., Corden, J. L., Westphal, H. and Beachy, P. A. (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* 383, 407-13.

Christian, J. L. and Moon, R. T. (1993). Interactions between Xwnt-8 and Spemann organizer signaling pathways generate dorsoventral pattern in the embryonic mesoderm of Xenopus. *Genes Dev* 7, 13-28.

Chun, S. Y., Chen, F., Washburn, J. G., MacDonald, J. W., Innes, K. L., Zhao, R., Cruz-Correa, M. R., Dang, L. H. and Dang, D. T. (2007). CDX2 promotes anchorage-independent growth by transcriptional repression of IGFBP-3. *Oncogene* 26, 4725-9.

Clevers, H. and Nusse, R. (2012). Wnt/beta-catenin signaling and disease. *Cell* 149, 1192-205.

Clevers, H. and van de Wetering, M. (1997). TCF/LEF factor earn their wings. *Trends Genet* 13, 485-9.

Colombres, M., Henriquez, J. P., Reig, G. F., Scheu, J., Calderon, R., Alvarez, A., Brandan, E. and Inestrosa, N. C. (2008). Heparin activates Wnt signaling for neuronal morphogenesis. *J Cell Physiol* 216, 805-15.

Condie, B. G. and Capecchi, M. R. (1993). Mice homozygous for a targeted disruption of Hoxd-3 (Hox-4.1) exhibit anterior transformations of the first and second cervical vertebrae, the atlas and the axis. *Development* 119, 579-95.

Conlon, R. A. and Rossant, J. (1992). Exogenous retinoic acid rapidly induces anterior ectopic expression of murine Hox-2 genes in vivo. *Development* 116, 357-68.

Constam, D. B. and Robertson, E. J. (1999). Regulation of bone morphogenetic protein activity by pro domains and proprotein convertases. *J Cell Biol* 144, 139-49.

Conway, S. J., Bundy, J., Chen, J., Dickman, E., Rogers, R. and Will, B. M. (2000). Decreased neural crest stem cell expansion is responsible for the conotruncal heart defects within the splotch (Sp(2H))/Pax3 mouse mutant. *Cardiovasc Res* 47, 314-28.

Copeland, N. G., Gilbert, D. J., Cho, B. C., Donovan, P. J., Jenkins, N. A., Cosman, D., Anderson, D., Lyman, S. D. and Williams, D. E. (1990). Mast cell growth factor maps near the steel locus on mouse chromosome 10 and is deleted in a number of steel alleles. *Cell* 63, 175-83.

Copf, T., Schroder, R. and Averof, M. (2004). Ancestral role of caudal genes in axis elongation and segmentation. *Proc Natl Acad Sci U S A* 101, 17711-5.

Copp, A. J. and Greene, N. D. (2013). Neural tube defects--disorders of neurulation and related embryonic processes. *Wiley Interdiscip Rev Dev Biol* 2, 213-27.

Copp, A. J., Greene, N. D. and Murdoch, J. N. (2003). The genetic basis of mammalian neurulation. *Nat Rev Genet* 4, 784-93.

Coucouvanis, E. and Martin, G. R. (1995). Signals for death and survival: a two-step mechanism for cavitation in the vertebrate embryo. *Cell* 83, 279-87.

Coutaud, B. and Pilon, N. (2013a). Characterization of a novel transgenic mouse line expressing Cre recombinase under the control of the Cdx2 neural specific enhancer. *Genesis* 51, 777-84.

Coutaud, B. and Pilon, N. (2013b). Characterization of a novel transgenic mouse line expressing Cre recombinase under the control of the Cdx2 neural specific enhancer. *Genesis*, 1-8.

Crissey, M. A., Guo, R. J., Funakoshi, S., Kong, J., Liu, J. and Lynch, J. P. (2011). Cdx2 levels modulate intestinal epithelium maturity and Paneth cell development. *Gastroenterology* 140, 517-528 e8.

Curran, K., Raible, D. W. and Lister, J. A. (2009). Foxd3 controls melanophore specification in the zebrafish neural crest by regulation of Mitf. *Dev Biol* 332, 408-17.

Daniels, D. L. and Weis, W. I. (2005). Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol* 12, 364-71.

David, S. G., Cebrian, C., Vaughan, E. D., Jr. and Herzlinger, D. (2005). c-kit and ureteral peristalsis. *J Urol* 173, 292-5.

Davidson, A. J., Ernst, P., Wang, Y., Dekens, M. P., Kingsley, P. D., Palis, J., Korsmeyer, S. J., Daley, G. Q. and Zon, L. I. (2003). cdx4 mutants fail to specify blood progenitors and can be rescued by multiple hox genes. *Nature* 425, 300-6.

Davidson, A. J. and Zon, L. I. (2006). The caudal-related homeobox genes cdx1a and cdx4 act redundantly to regulate hox gene expression and the formation of putative hematopoietic stem cells during zebrafish embryogenesis. *Dev Biol* 292, 506-18.

de Croze, N., Maczkowiak, F. and Monsoro-Burq, A. H. (2011). Reiterative AP2a activity controls sequential steps in the neural crest gene regulatory network. *Proc Natl Acad Sci U S A* 108, 155-60.

De Robertis, E. M. and Kuroda, H. (2004). Dorsal-ventral patterning and neural induction in Xenopus embryos. *Annu Rev Cell Dev Biol* 20, 285-308.

De Robertis, E. M. and Sasai, Y. (1996). A common plan for dorsoventral patterning in Bilateria. *Nature* 380, 37-40.

de Sousa Lopes, S. M., Carvalho, R. L., van den Driesche, S., Goumans, M. J., ten Dijke, P. and Mummery, C. L. (2003). Distribution of phosphorylated Smad2 identifies target tissues of TGF beta ligands in mouse development. *Gene Expr Patterns* 3, 355-60.

Dearolf, C. R., Topol, J. and Parker, C. S. (1989). The caudal gene product is a direct activator of fushi tarazu transcription during Drosophila embryogenesis. *Nature* 341, 340-3.

Degenhardt, K. R., Milewski, R. C., Padmanabhan, A., Miller, M., Singh, M. K., Lang, D., Engleka, K. A., Wu, M., Li, J., Zhou, D. et al. (2010). Distinct enhancers at the Pax3 locus can function redundantly to regulate neural tube and neural crest expressions. *Dev Biol* 339, 519-27.

del Barco Barrantes, I., Davidson, G., Grone, H. J., Westphal, H. and Niehrs, C. (2003). Dkk1 and noggin cooperate in mammalian head induction. *Genes Dev* 17, 2239-44.

Deschamps, J., van den Akker, E., Forlani, S., De Graaff, W., Oosterveen, T., Roelen, B. and Roelfsema, J. (1999). Initiation, establishment and maintenance of Hox gene expression patterns in the mouse. *Int J Dev Biol* 43, 635-50.

Deschamps, J. and van Nes, J. (2005). Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* 132, 2931-42.

Dessaud, E., McMahon, A. P. and Briscoe, J. (2008). Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network. *Development* 135, 2489-503.

Di-Gregorio, A., Sancho, M., Stuckey, D. W., Crompton, L. A., Godwin, J., Mishina, Y. and Rodriguez, T. A. (2007). BMP signalling inhibits premature neural differentiation in the mouse embryo. *Development* 134, 3359-69.

Domon-Dell, C. and Freund, J. N. (2002). Stimulation of Cdx1 by oncogenic beta-catenin/Tcf4 in colon cancer cells; opposite effect of the CDX2 homeoprotein. *FEBS Lett* 518, 83-7.

Dottori, M., Gross, M. K., Labosky, P. and Goulding, M. (2001). The winged-helix transcription factor Foxd3 suppresses interneuron differentiation and promotes neural crest cell fate. *Development* 128, 4127-38.

Drummond, F., Putt, W., Fox, M. and Edwards, Y. H. (1997). Cloning and chromosome assignment of the human CDX2 gene. *Ann Hum Genet* 61, 393-400.

Dubrulle, J. and Pourquie, O. (2004). Coupling segmentation to axis formation. *Development* 131, 5783-93.

Dudley, J. A., Haworth, J. M., McGraw, M. E., Frank, J. D. and Tizard, E. J. (1997). Clinical relevance and implications of antenatal hydronephrosis. *Arch Dis Child Fetal Neonatal Ed* 76, F31-4.

Duprey, P., Chowdhury, K., Dressler, G. R., Balling, R., Simon, D., Guenet, J. L. and Gruss, P. (1988). A mouse gene homologous to the Drosophila gene caudal is expressed in epithelial cells from the embryonic intestine. *Genes Dev* 2, 1647-54.

Duval, N., Daubas, P., Bourcier de Carbon, C., St Cloment, C., Tinevez, J. Y., Lopes, M., Ribes, V. and Robert, B. (2014). Msx1 and Msx2 act as essential activators of Atoh1 expression in the murine spinal cord. *Development* 141, 1726-36.

Ebert, P. J., Timmer, J. R., Nakada, Y., Helms, A. W., Parab, P. B., Liu, Y., Hunsaker, T. L. and Johnson, J. E. (2003). Zic1 represses Math1 expression via interactions with the Math1 enhancer and modulation of Math1 autoregulation. *Development* 130, 1949-59.

Echelard, Y., Epstein, D. J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. A. and McMahon, A. P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* 75, 1417-30.

Elms, P., Siggers, P., Napper, D., Greenfield, A. and Arkell, R. (2003). Zic2 is required for neural crest formation and hindbrain patterning during mouse development. *Dev Biol* 264, 391-406.

Enomoto, H., Araki, T., Jackman, A., Heuckeroth, R. O., Snider, W. D., Johnson, E. M., Jr. and Milbrandt, J. (1998). GFR alpha1-deficient mice have deficits in the enteric nervous system and kidneys. *Neuron* 21, 317-24.

Epperlein, H. H., Selleck, M. A., Meulemans, D., McHedlishvili, L., Cerny, R., Sobkow, L. and Bronner-Fraser, M. (2007). Migratory patterns and developmental potential of trunk neural crest cells in the axolotl embryo. *Dev Dyn* 236, 389-403.

Epstein, D. J., Vekemans, M. and Gros, P. (1991). Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell* 67, 767-74.

Epstein, D. J., Vogan, K. J., Trasler, D. G. and Gros, P. (1993). A mutation within intron 3 of the Pax-3 gene produces aberrantly spliced mRNA transcripts in the splotch (Sp) mouse mutant. *Proc Natl Acad Sci U S A* 90, 532-6.

Epstein, J. A., Li, J., Lang, D., Chen, F., Brown, C. B., Jin, F., Lu, M. M., Thomas, M., Liu, E., Wessels, A. et al. (2000). Migration of cardiac neural crest cells in Splotch embryos. *Development* 127, 1869-78.

Erickson, C. A. and Weston, J. A. (1983). An SEM analysis of neural crest migration in the mouse. *J Embryol Exp Morphol* 74, 97-118.

Ericson, J., Morton, S., Kawakami, A., Roelink, H. and Jessell, T. M. (1996). Two critical periods of Sonic Hedgehog signaling required for the specification of motor neuron identity. *Cell* 87, 661-73.

Estibeiro, J. P., Brook, F. A. and Copp, A. J. (1993). Interaction between splotch (Sp) and curly tail (ct) mouse mutants in the embryonic development of neural tube defects. *Development* 119, 113-21.

Faas, L. and Isaacs, H. V. (2009). Overlapping functions of Cdx1, Cdx2, and Cdx4 in the development of the amphibian Xenopus tropicalis. *Dev Dyn* 238, 835-52.

Fainsod, A., Deissler, K., Yelin, R., Marom, K., Epstein, M., Pillemer, G., Steinbeisser, H. and Blum, M. (1997). The dorsalizing and neural inducing gene follistatin is an antagonist of BMP-4. *Mech Dev* 63, 39-50.

Finnerty, J. R., Mazza, M. E. and Jezewski, P. A. (2009). Domain duplication, divergence, and loss events in vertebrate Msx paralogs reveal phylogenomically informed disease markers. *BMC Evol Biol* 9, 18.

Fossat, N., Jones, V., Garcia-Garcia, M. J. and Tam, P. P. (2012). Modulation of WNT signaling activity is key to the formation of the embryonic head. *Cell Cycle* 11, 26-32.

Franz, T. (1992). Neural tube defects without neural crest defects in splotch mice. *Teratology* 46, 599-604.

Fuentealba, L. C., Eivers, E., Ikeda, A., Hurtado, C., Kuroda, H., Pera, E. M. and De Robertis, E. M. (2007). Integrating patterning signals: Wnt/GSK3 regulates the duration of the BMP/Smad1 signal. *Cell* 131, 980-93.

Gamer, L. W. and Wright, C. V. (1993). Murine Cdx-4 bears striking similarities to the Drosophila caudal gene in its homeodomain sequence and early expression pattern. *Mech Dev* 43, 71-81.

Gammill, L. S. and Bronner-Fraser, M. (2003). Neural crest specification: migrating into genomics. *Nat Rev Neurosci* 4, 795-805.

Gammill, L. S., Gonzalez, C., Gu, C. and Bronner-Fraser, M. (2006). Guidance of trunk neural crest migration requires neuropilin 2/semaphorin 3F signaling. *Development* 133, 99-106.

Gans, C. and Northcutt, R. G. (1983). Neural crest and the origin of vertebrates: a new head. *Science* 220, 268-73.

Gao, N., White, P. and Kaestner, K. H. (2009). Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. *Dev Cell* 16, 588-99.

Garcia-Castro, M. I., Marcelle, C. and Bronner-Fraser, M. (2002). Ectodermal Wnt function as a neural crest inducer. *Science* 297, 848-51.

Garcia-Garcia, M. J. and Anderson, K. V. (2003). Essential role of glycosaminoglycans in Fgf signaling during mouse gastrulation. *Cell* 114, 727-37.

Gardner, R. L. (1978). The relationship between cell lineage and differentiation in the early mouse embryo. *Results Probl Cell Differ* 9, 205-41.

Gardner, R. L. (1983). Origin and differentiation of extraembryonic tissues in the mouse. *Int Rev Exp Pathol* 24, 63-133.

Garnett, A. T., Square, T. A. and Medeiros, D. M. (2012). BMP, Wnt and FGF signals are integrated through evolutionarily conserved enhancers to achieve robust expression of Pax3 and Zic genes at the zebrafish neural plate border. *Development* 139, 4220-31.

Gaston-Massuet, C., Henderson, D. J., Greene, N. D. and Copp, A. J. (2005). Zic4, a zinc-finger transcription factor, is expressed in the developing mouse nervous system. *Dev Dyn* 233, 1110-5.

Gaunt, S. J., Cockley, A. and Drage, D. (2004). Additional enhancer copies, with intact cdx binding sites, anteriorize Hoxa-7/lacZ expression in mouse embryos: evidence in keeping with an instructional cdx gradient. *Int J Dev Biol* 48, 613-22.

Gaunt, S. J., Drage, D. and Cockley, A. (2003). Vertebrate caudal gene expression gradients investigated by use of chick cdx-A/lacZ and mouse cdx-1/lacZ reporters in transgenic mouse embryos: evidence for an intron enhancer. *Mech Dev* 120, 573-86.

Gaunt, S. J., Drage, D. and Trubshaw, R. C. (2005). cdx4/lacZ and cdx2/lacZ protein gradients formed by decay during gastrulation in the mouse. *Int J Dev Biol* 49, 901-8.

Gaunt, S. J., Drage, D. and Trubshaw, R. C. (2008). Increased Cdx protein dose effects upon axial patterning in transgenic lines of mice. *Development* 135, 2511-20.

Gaunt, S. J. and Paul, Y. L. (2011). Origins of Cdx1 regulatory elements suggest roles in vertebrate evolution. *Int J Dev Biol* 55, 93-8.

Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994). Homeodomain-DNA recognition. *Cell* 78, 211-23.

Geissler, E. N., Ryan, M. A. and Housman, D. E. (1988). The dominant-white spotting (W) locus of the mouse encodes the c-kit proto-oncogene. *Cell* 55, 185-92.

German, M. S., Wang, J., Fernald, A. A., Espinosa, R., 3rd, Le Beau, M. M. and Bell, G. I. (1994). Localization of the genes encoding two transcription factors, LMX1 and CDX3, regulating insulin gene expression to human chromosomes 1 and 13. *Genomics* 24, 403-4.

Glinka, A., Wu, W., Delius, H., Monaghan, A. P., Blumenstock, C. and Niehrs, C. (1998). Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature* 391, 357-62.

Glinka, A., Wu, W., Onichtchouk, D., Blumenstock, C. and Niehrs, C. (1997). Head induction by simultaneous repression of Bmp and Wnt signalling in Xenopus. *Nature* 389, 517-9.

Gofflot, F., Hall, M. and Morriss-Kay, G. M. (1997). Genetic patterning of the developing mouse tail at the time of posterior neuropore closure. *Dev Dyn* 210, 431-45.

Gong, Y., Bourhis, E., Chiu, C., Stawicki, S., DeAlmeida, V. I., Liu, B. Y., Phamluong, K., Cao, T. C., Carano, R. A., Ernst, J. A. et al. (2010). Wnt isoform-specific interactions with coreceptor specify inhibition or potentiation of signaling by LRP6 antibodies. *PLoS One* 5, e12682.

Goodrich, L. V., Milenkovic, L., Higgins, K. M. and Scott, M. P. (1997). Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science* 277, 1109-13.

Gordon, M. D. and Nusse, R. (2006). Wnt signaling: multiple pathways, multiple receptors, and multiple transcription factors. *J Biol Chem* 281, 22429-33.

Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. R. and Gruss, P. (1991). Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J* 10, 1135-47.

Goulding, M. D., Lumsden, A. and Gruss, P. (1993). Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord. *Development* 117, 1001-16.

Grainger, S., Hryniuk, A. and Lohnes, D. (2013). Cdx1 and Cdx2 exhibit transcriptional specificity in the intestine. *PLoS One* 8, e54757.

Grainger, S., Lam, J., Savory, J. G., Mears, A. J., Rijli, F. M. and Lohnes, D. (2011). Cdx regulates Dll1 in multiple lineages. *Dev Biol*.

Grainger, S., Lam, J., Savory, J. G., Mears, A. J., Rijli, F. M. and Lohnes, D. (2012). Cdx regulates Dll1 in multiple lineages. *Dev Biol* 361, 1-11.

Grainger, S., Savory, J. G. and Lohnes, D. (2010). Cdx2 regulates patterning of the intestinal epithelium. *Dev Biol* 339, 155-65.

Green, S. A. and Bronner, M. E. (2013). Gene duplications and the early evolution of neural crest development. *Semin Cell Dev Biol* 24, 95-100.

Greene, N. D. and Copp, A. J. (2014). Neural tube defects. *Annu Rev Neurosci* 37, 221-42.

Greene, N. D., Stanier, P. and Copp, A. J. (2009). Genetics of human neural tube defects. *Hum Mol Genet* 18, R113-29.

Grinberg, I., Northrup, H., Ardinger, H., Prasad, C., Dobyns, W. B. and Millen, K. J. (2004). Heterozygous deletion of the linked genes ZIC1 and ZIC4 is involved in Dandy-Walker malformation. *Nat Genet* 36, 1053-5.

Groppe, J., Greenwald, J., Wiater, E., Rodriguez-Leon, J., Economides, A. N., Kwiatkowski, W., Affolter, M., Vale, W. W., Izpisua Belmonte, J. C. and Choe, S. (2002). Structural basis of BMP signalling inhibition by the cystine knot protein Noggin. *Nature* 420, 636-42.

Grumolato, L., Liu, G., Mong, P., Mudbhary, R., Biswas, R., Arroyave, R., Vijayakumar, S., Economides, A. N. and Aaronson, S. A. (2010). Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. *Genes Dev* 24, 2517-30.

Grunz, H. and Tacke, L. (1989). Neural differentiation of Xenopus laevis ectoderm takes place after disaggregation and delayed reaggregation without inducer. *Cell Differ Dev* 28, 211-7.

Guo, R. J., Funakoshi, S., Lee, H. H., Kong, J. and Lynch, J. P. (2010). The intestine-specific transcription factor Cdx2 inhibits beta-catenin/TCF transcriptional activity by disrupting the beta-catenin-TCF protein complex. *Carcinogenesis* 31, 159-66.

Guyonneau, L., Murisier, F., Rossier, A., Moulin, A. and Beermann, F. (2004). Melanocytes and pigmentation are affected in dopachrome tautomerase knockout mice. *Mol Cell Biol* 24, 3396-403.

Hall, B. K. (2000). The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. *Evol Dev* 2, 3-5.

Hallikas, O., Palin, K., Sinjushina, N., Rautiainen, R., Partanen, J., Ukkonen, E. and Taipale, J. (2006). Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* 124, 47-59.

Hammerschmidt, M., Serbedzija, G. N. and McMahon, A. P. (1996). Genetic analysis of dorsoventral pattern formation in the zebrafish: requirement of a BMP-like ventralizing activity and its dorsal repressor. *Genes Dev* 10, 2452-61.

Hanna, L. A., Foreman, R. K., Tarasenko, I. A., Kessler, D. S. and Labosky, P. A. (2002). Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. *Genes Dev* 16, 2650-61.

Hari, L., Miescher, I., Shakhova, O., Suter, U., Chin, L., Taketo, M., Richardson, W. D., Kessaris, N. and Sommer, L. (2012). Temporal control of neural crest lineage generation by Wnt/beta-catenin signaling. *Development* 139, 2107-17.

Harland, R. and Gerhart, J. (1997). Formation and function of Spemann's organizer. *Annu Rev Cell Dev Biol* 13, 611-67.

Harris, M. J. and Juriloff, D. M. (2010). An update to the list of mouse mutants with neural tube closure defects and advances toward a complete genetic perspective of neural tube closure. *Birth Defects Res A Clin Mol Teratol* 88, 653-69.

Hayashi, K., Kobayashi, T., Umino, T., Goitsuka, R., Matsui, Y. and Kitamura, D. (2002). SMAD1 signaling is critical for initial commitment of germ cell lineage from mouse epiblast. *Mech Dev* 118, 99-109.

He, S., Pant, D., Schiffmacher, A., Meece, A. and Keefer, C. L. (2008). Lymphoid enhancer factor 1-mediated Wnt signaling promotes the initiation of trophoblast lineage differentiation in mouse embryonic stem cells. *Stem Cells* 26, 842-9.

He, S. J., Stevens, G., Braithwaite, A. W. and Eccles, M. R. (2005). Transfection of melanoma cells with antisense PAX3 oligonucleotides additively complements cisplatin-induced cytotoxicity. *Mol Cancer Ther* 4, 996-1003.

He, X., Treacy, M. N., Simmons, D. M., Ingraham, H. A., Swanson, L. W. and Rosenfeld, M. G. (1989). Expression of a large family of POU-domain regulatory genes in mammalian brain development. *Nature* 340, 35-41.

Heimbucher, T., Murko, C., Bajoghli, B., Aghaallaei, N., Huber, A., Stebegg, R., Eberhard, D., Fink, M., Simeone, A. and Czerny, T. (2007). Gbx2 and Otx2 interact with the WD40 domain of Groucho/Tle corepressors. *Mol Cell Biol* 27, 340-51.

Hemmati-Brivanlou, A., Kelly, O. G. and Melton, D. A. (1994). Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell* 77, 283-95.

Hemmati-Brivanlou, A. and Melton, D. (1997). Vertebrate embryonic cells will become nerve cells unless told otherwise. *Cell* 88, 13-7.

Herpin, A. and Cunningham, C. (2007). Cross-talk between the bone morphogenetic protein pathway and other major signaling pathways results in tightly regulated cell-specific outcomes. *FEBS J* 274, 2977-85.

Heydemann, A., Nguyen, L. C. and Crenshaw, E. B., 3rd. (2001). Regulatory regions from the Brn4 promoter direct LACZ expression to the developing forebrain and neural tube. *Brain Res Dev Brain Res* 128, 83-90.

Hikasa, H., Ezan, J., Itoh, K., Li, X., Klymkowsky, M. W. and Sokol, S. Y. (2010). Regulation of TCF3 by Wnt-dependent phosphorylation during vertebrate axis specification. *Dev Cell* 19, 521-32.

Hikasa, H. and Sokol, S. Y. (2013). Wnt signaling in vertebrate axis specification. *Cold Spring Harb Perspect Biol* 5, a007955.

His, W. (1868). Die erste entwickelung des huhnchens im Ei. Leipzig: FCW Vogel. *Untersuchungen uber die erste Anlage des Wirbeltierleibes*.Hoffman, B. G., Robertson, G., Zavaglia, B., Beach, M., Cullum, R., Lee, S., Soukhatcheva, G., Li, L., Wederell, E. D., Thiessen, N. et al. (2010). Locus co-occupancy, nucleosome positioning, and H3K4me1 regulate the functionality of FOXA2-, HNF4A-, and PDX1-bound loci in islets and liver. *Genome Res* 20, 1037-51.

Hogan, B. L. (1996). Bone morphogenetic proteins: multifunctional regulators of vertebrate development. *Genes Dev* 10, 1580-94.

Hogan, B. L. (1999). Morphogenesis. *Cell* 96, 225-33.

Holland, L. Z. (2009). Chordate roots of the vertebrate nervous system: expanding the molecular toolkit. *Nat Rev Neurosci* 10, 736-46.

Holland, L. Z., Schubert, M., Kozmik, Z. and Holland, N. D. (1999). AmphiPax3/7, an amphioxus paired box gene: insights into chordate myogenesis, neurogenesis, and

the possible evolutionary precursor of definitive vertebrate neural crest. *Evol Dev* 1, 153-65.

Holland, P. W. (1991). Cloning and evolutionary analysis of msh-like homeobox genes from mouse, zebrafish and ascidian. *Gene* 98, 253-7.

Holtz, A. M., Peterson, K. A., Nishi, Y., Morin, S., Song, J. Y., Charron, F., McMahon, A. P. and Allen, B. L. (2013). Essential role for ligand-dependent feedback antagonism of vertebrate hedgehog signaling by PTCH1, PTCH2 and HHIP1 during neural patterning. *Development* 140, 3423-34.

Hooper, J. E. and Scott, M. P. (1989). The Drosophila patched gene encodes a putative membrane protein required for segmental patterning. *Cell* 59, 751-65.

Hoppler, S., Brown, J. D. and Moon, R. T. (1996). Expression of a dominant-negative Wnt blocks induction of MyoD in Xenopus embryos. *Genes Dev* 10, 2805-17.

Horn, J. M. and Ashworth, A. (1995). A member of the caudal family of homeobox genes maps to the X-inactivation centre region of the mouse and human X chromosomes. *Hum Mol Genet* 4, 1041-7.

Hornyak, T. J., Hayes, D. J., Chiu, L. Y. and Ziff, E. B. (2001). Transcription factors in melanocyte development: distinct roles for Pax-3 and Mitf. *Mech Dev* 101, 47-59.

Hosoda, K., Hammer, R. E., Richardson, J. A., Baynash, A. G., Cheung, J. C., Giaid, A. and Yanagisawa, M. (1994). Targeted and natural (piebald-lethal) mutations of endothelin-B receptor gene produce megacolon associated with spotted coat color in mice. *Cell* 79, 1267-76.

Houle, M., Prinos, P., Iulianella, A., Bouchard, N. and Lohnes, D. (2000). Retinoic acid regulation of Cdx1: an indirect mechanism for retinoids and vertebral specification. *Mol Cell Biol* 20, 6579-86.

Houle, M., Sylvestre, J. R. and Lohnes, D. (2003). Retinoic acid regulates a subset of Cdx1 function in vivo. *Development* 130, 6555-67.

Houtmeyers, R., Souopgui, J., Tejpar, S. and Arkell, R. (2013). The ZIC gene family encodes multi-functional proteins essential for patterning and morphogenesis. *Cell Mol Life Sci* 70, 3791-811.

Houzelstein, D., Cohen, A., Buckingham, M. E. and Robert, B. (1997). Insertional mutation of the mouse Msx1 homeobox gene by an nlacZ reporter gene. *Mech Dev* 65, 123-33.

Hryniuk, A., Grainger, S., Savory, J. G. and Lohnes, D. (2014). Cdx1 and cdx2 function as tumor suppressors. *J Biol Chem* 289, 33343-54.

Huang, B. L., Brugger, S. M. and Lyons, K. M. (2010). Stage-specific control of connective tissue growth factor (CTGF/CCN2) expression in chondrocytes by Sox9 and beta-catenin. *J Biol Chem* 285, 27702-12.

Hui, C. C. and Angers, S. (2011). Gli proteins in development and disease. *Annu Rev Cell Dev Biol* 27, 513-37.

Hurtado, R., Bub, G. and Herzlinger, D. (2010). The pelvis-kidney junction contains HCN3, a hyperpolarization-activated cation channel that triggers ureter peristalsis. *Kidney Int* 77, 500-8.

Hussain, M. A. and Habener, J. F. (1999). Glucagon gene transcription activation mediated by synergistic interactions of pax-6 and cdx-2 with the p300 co-activator. *J Biol Chem* 274, 28950-7.

Iimura, T. and Pourquie, O. (2006). Collinear activation of Hoxb genes during gastrulation is linked to mesoderm cell ingression. *Nature* 442, 568-71.

Ikeya, M. and Takada, S. (2001). Wnt-3a is required for somite specification along the anteroposterior axis of the mouse embryo and for regulation of cdx-1 expression. *Mech Dev* 103, 27-33.

Imamura, T., Takase, M., Nishihara, A., Oeda, E., Hanai, J., Kawabata, M. and Miyazono, K. (1997). Smad6 inhibits signalling by the TGF-beta superfamily. *Nature* 389, 622-6.

Incardona, J. P., Gruenberg, J. and Roelink, H. (2002). Sonic hedgehog induces the segregation of patched and smoothened in endosomes. *Curr Biol* 12, 983-95.

Ingham, P. W. and McMahon, A. P. (2001). Hedgehog signaling in animal development: paradigms and principles. *Genes Dev* 15, 3059-87.

Inoue, T., Hatayama, M., Tohmonda, T., Itohara, S., Aruga, J. and Mikoshiba, K. (2004). Mouse Zic5 deficiency results in neural tube defects and hypoplasia of cephalic neural crest derivatives. *Dev Biol* 270, 146-62.

Inoue, T., Ogawa, M., Mikoshiba, K. and Aruga, J. (2008). Zic deficiency in the cortical marginal zone and meninges results in cortical lamination defects resembling those in type II lissencephaly. *J Neurosci* 28, 4712-25.

Inoue, T., Ota, M., Mikoshiba, K. and Aruga, J. (2007a). Zic2 and Zic3 synergistically control neurulation and segmentation of paraxial mesoderm in mouse embryo. *Dev Biol* 306, 669-84.

Inoue, T., Ota, M., Ogawa, M., Mikoshiba, K. and Aruga, J. (2007b). Zic1 and Zic3 regulate medial forebrain development through expansion of neuronal progenitors. *J Neurosci* 27, 5461-73.

Isaacs, H. V., Pownall, M. E. and Slack, J. M. (1998). Regulation of Hox gene expression and posterior development by the Xenopus caudal homologue Xcad3. *EMBO J* 17, 3413-27.

Ishii, M., Han, J., Yen, H. Y., Sucov, H. M., Chai, Y. and Maxson, R. E., Jr. (2005). Combined deficiencies of Msx1 and Msx2 cause impaired patterning and survival of the cranial neural crest. *Development* 132, 4937-50.

James, R. and Kazenwadel, J. (1991). Homeobox gene expression in the intestinal epithelium of adult mice. *J Biol Chem* 266, 3246-51.

Jandzik, D., Garnett, A. T., Square, T. A., Cattell, M. V., Yu, J. K. and Medeiros, D. M. (2014). Evolution of the new vertebrate head by co-option of an ancient chordate skeletal tissue. *Nature*.

Jarad, G. and Miner, J. H. (2009). The Pax3-Cre transgene exhibits a rostrocaudal gradient of expression in the skeletal muscle lineage. *Genesis* 47, 1-6.

Jaynes, J. B. and O'Farrell, P. H. (1991). Active repression of transcription by the engrailed homeodomain protein. *EMBO J* 10, 1427-33.

Jeffery, W. R. (2006). Ascidian neural crest-like cells: phylogenetic distribution, relationship to larval complexity, and pigment cell fate. *J Exp Zool B Mol Dev Evol* 306, 470-80.

Jeffery, W. R., Chiba, T., Krajka, F. R., Deyts, C., Satoh, N. and Joly, J. S. (2008). Trunk lateral cells are neural crest-like cells in the ascidian Ciona intestinalis: insights into the ancestry and evolution of the neural crest. *Dev Biol* 324, 152-60.

Jeffery, W. R., Strickler, A. G. and Yamamoto, Y. (2004). Migratory neural crest-like cells form body pigmentation in a urochordate embryo. *Nature* 431, 696-9.

Jeong, J. and McMahon, A. P. (2005). Growth and pattern of the mammalian neural tube are governed by partially overlapping feedback activities of the hedgehog antagonists patched 1 and Hhip1. *Development* 132, 143-54.

Jessell, T. M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat Rev Genet* 1, 20-9.

Jho, E. H., Zhang, T., Domon, C., Joo, C. K., Freund, J. N. and Costantini, F. (2002). Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Mol Cell Biol* 22, 1172-83.

Johnson, M. H. and Ziomek, C. A. (1981). The foundation of two distinct cell lineages within the mouse morula. *Cell* 24, 71-80.

Jones-Villeneuve, E. M., Rudnicki, M. A., Harris, J. F. and McBurney, M. W. (1983). Retinoic acid-induced neural differentiation of embryonal carcinoma cells. *Mol Cell Biol* 3, 2271-9.

Jonk, L. J., Itoh, S., Heldin, C. H., ten Dijke, P. and Kruijer, W. (1998). Identification and functional characterization of a Smad binding element (SBE) in the JunB promoter that acts as a transforming growth factor-beta, activin, and bone morphogenetic protein-inducible enhancer. *J Biol Chem* 273, 21145-52.

Joo, J. H., Taxter, T. J., Munguba, G. C., Kim, Y. H., Dhaduvai, K., Dunn, N. W., Degan, W. J., Oh, S. P. and Sugrue, S. P. (2010). Pinin modulates expression of an intestinal homeobox gene, Cdx2, and plays an essential role for small intestinal morphogenesis. *Dev Biol* 345, 191-203.

Jumlongras, D., Bei, M., Stimson, J. M., Wang, W. F., DePalma, S. R., Seidman, C. E., Felbor, U., Maas, R., Seidman, J. G. and Olsen, B. R. (2001). A nonsense mutation in MSX1 causes Witkop syndrome. *Am J Hum Genet* 69, 67-74.

Kam, M. K., Cheung, M., Zhu, J. J., Cheng, W. W., Sat, E. W., Tam, P. K. and Lui, V. C. (2014a). Homeobox b5 (Hoxb5) regulates the expression of Forkhead box D3 gene (Foxd3) in neural crest. *Int J Biochem Cell Biol* 55C, 144-152.

Kam, M. K., Cheung, M. C., Zhu, J. J., Cheng, W. W., Sat, E. W., Tam, P. K. and Lui, V. C. (2014b). Perturbation of Hoxb5 signaling in vagal and trunk neural crest cells causes apoptosis and neurocristopathies in mice. *Cell Death Differ* 21, 278-89.

Kam, Y. W., Okumura, Y., Kido, H., Ng, L. F., Bruzzone, R. and Altmeyer, R. (2009). Cleavage of the SARS coronavirus spike glycoprotein by airway proteases enhances virus entry into human bronchial epithelial cells in vitro. *PLoS One* 4, e7870.

Kaneko, K. J., Kohn, M. J., Liu, C. and DePamphilis, M. L. (2007). Transcription factor TEAD2 is involved in neural tube closure. *Genesis* 45, 577-87.

Katsuyama, Y., Sato, Y., Wada, S. and Saiga, H. (1999). Ascidian tail formation requires caudal function. *Dev Biol* 213, 257-68.

Kaufman, M. H. (1992). The Atlas of Mouse Development. London: Academic Press.

Keenan, I. D., Sharrard, R. M. and Isaacs, H. V. (2006). FGF signal transduction and the regulation of Cdx gene expression. *Dev Biol* 299, 478-88.

Keller, R., Shook, D. and Skoglund, P. (2008). The forces that shape embryos: physical aspects of convergent extension by cell intercalation. *Phys Biol* 5, 015007.

Kelly, O. G., Pinson, K. I. and Skarnes, W. C. (2004). The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice. *Development* 131, 2803-15.

Kessel, M. and Gruss, P. (1991). Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* 67, 89-104.

Kibar, Z., Vogan, K. J., Groulx, N., Justice, M. J., Underhill, D. A. and Gros, P. (2001). Ltap, a mammalian homolog of Drosophila Strabismus/Van Gogh, is altered in the mouse neural tube mutant Loop-tail. *Nat Genet* 28, 251-5.

Kieny, M., Mauger, A. and Sengel, P. (1972). Early regionalization of somitic mesoderm as studied by the development of axial skeleton of the chick embryo. *Dev Biol* 28, 142-61.

Kim, C. H., Oda, T., Itoh, M., Jiang, D., Artinger, K. B., Chandrasekharappa, S. C., Driever, W. and Chitnis, A. B. (2000). Repressor activity of Headless/Tcf3 is essential for vertebrate head formation. *Nature* 407, 913-6.

Kinder, S. J., Tsang, T. E., Wakamiya, M., Sasaki, H., Behringer, R. R., Nagy, A. and Tam, P. P. (2001). The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial mesoderm. *Development* 128, 3623-34.

Kinzler, K. W., Nilbert, M. C., Su, L. K., Vogelstein, B., Bryan, T. M., Levy, D. B., Smith, K. J., Preisinger, A. C., Hedge, P., McKechnie, D. et al. (1991). Identification of FAP locus genes from chromosome 5q21. *Science* 253, 661-5.

Komiya, Y. and Habas, R. (2008). Wnt signal transduction pathways. *Organogenesis* 4, 68-75.

Kormish, J. D., Sinner, D. and Zorn, A. M. (2010). Interactions between SOX factors and Wnt/beta-catenin signaling in development and disease. *Dev Dyn* 239, 56-68.

Kos, R., Reedy, M. V., Johnson, R. L. and Erickson, C. A. (2001). The winged-helix transcription factor FoxD3 is important for establishing the neural crest lineage and repressing melanogenesis in avian embryos. *Development* 128, 1467-79.

Kothary, R., Clapoff, S., Darling, S., Perry, M. D., Moran, L. A. and Rossant, J. (1989). Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. *Development* 105, 707-14.

Koyabu, Y., Nakata, K., Mizugishi, K., Aruga, J. and Mikoshiba, K. (2001). Physical and functional interactions between Zic and Gli proteins. *J Biol Chem* 276, 6889-92.

Krauss, S., Concordet, J. P. and Ingham, P. W. (1993). A functionally conserved homolog of the Drosophila segment polarity gene hh is expressed in tissues with polarizing activity in zebrafish embryos. *Cell* 75, 1431-44.

Krispin, S., Nitzan, E., Kassem, Y. and Kalcheim, C. (2010). Evidence for a dynamic spatiotemporal fate map and early fate restrictions of premigratory avian neural crest. *Development* 137, 585-95.

Krotoski, D. M., Fraser, S. E. and Bronner-Fraser, M. (1988). Mapping of neural crest pathways in Xenopus laevis using inter- and intra-specific cell markers. *Dev Biol* 127, 119-32.

Krumlauf, R. (1994). Hox genes in vertebrate development. *Cell* 78, 191-201.

Kubic, J. D., Young, K. P., Plummer, R. S., Ludvik, A. E. and Lang, D. (2008). Pigmentation PAX-ways: the role of Pax3 in melanogenesis, melanocyte stem cell maintenance, and disease. *Pigment Cell Melanoma Res* 21, 627-45.

Kuwabara, T., Hsieh, J., Muotri, A., Yeo, G., Warashina, M., Lie, D. C., Moore, L., Nakashima, K., Asashima, M. and Gage, F. H. (2009). Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. *Nat Neurosci* 12, 1097-105.

Kuzuoka, M., Takahashi, T., Guron, C. and Raghow, R. (1994). Murine homeobox-containing gene, Msx-1: analysis of genomic organization, promoter structure, and potential autoregulatory cis-acting elements. *Genomics* 21, 85-91.

Kwang, S. J., Brugger, S. M., Lazik, A., Merrill, A. E., Wu, L. Y., Liu, Y. H., Ishii, M., Sangiorgi, F. O., Rauchman, M., Sucov, H. M. et al. (2002). Msx2 is an immediate downstream effector of Pax3 in the development of the murine cardiac neural crest. *Development* 129, 527-38.

LaBonne, C. and Bronner-Fraser, M. (1998). Neural crest induction in Xenopus: evidence for a two-signal model. *Development* 125, 2403-14.

Labosky, P. A. and Kaestner, K. H. (1998). The winged helix transcription factor Hfh2 is expressed in neural crest and spinal cord during mouse development. *Mech Dev* 76, 185-90.

Lafontaine, C. A., Grainger, S., Hess, B. L., Beland, M. and Lohnes, D. (2012). Cdx1 interacts physically with a subset of Hox proteins. *Biochemistry* 51, 9698-705.

Lallemand, Y., Nicola, M. A., Ramos, C., Bach, A., Clement, C. S. and Robert, B. (2005). Analysis of Msx1; Msx2 double mutants reveals multiple roles for Msx genes in limb development. *Development* 132, 3003-14.

Lamb, T. D., Collin, S. P. and Pugh, E. N., Jr. (2007). Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. *Nat Rev Neurosci* 8, 960-76.

Lamb, T. M. and Harland, R. M. (1995). Fibroblast growth factor is a direct neural inducer, which combined with noggin generates anterior-posterior neural pattern. *Development* 121, 3627-36.

Lamb, T. M., Knecht, A. K., Smith, W. C., Stachel, S. E., Economides, A. N., Stahl, N., Yancopolous, G. D. and Harland, R. M. (1993). Neural induction by the secreted polypeptide noggin. *Science* 262, 713-8.

Lane, P. W. and Liu, H. M. (1984). Association of megacolon with a new dominant spotting gene (Dom) in the mouse. *J Hered* 75, 435-9.

Lang, D., Chen, F., Milewski, R., Li, J., Lu, M. M. and Epstein, J. A. (2000). Pax3 is required for enteric ganglia formation and functions with Sox10 to modulate expression of c-ret. *J Clin Invest* 106, 963-71.

Lang, D., Lu, M. M., Huang, L., Engleka, K. A., Zhang, M., Chu, E. Y., Lipner, S., Skoultchi, A., Millar, S. E. and Epstein, J. A. (2005). Pax3 functions at a nodal point in melanocyte stem cell differentiation. *Nature* 433, 884-7.

Lang, D., Mascarenhas, J. B. and Shea, C. R. (2013). Melanocytes, melanocyte stem cells, and melanoma stem cells. *Clin Dermatol* 31, 166-78.

Lang, D., Powell, S. K., Plummer, R. S., Young, K. P. and Ruggeri, B. A. (2007). PAX genes: roles in development, pathophysiology, and cancer. *Biochem Pharmacol* 73, 1-14.

Lang, R. J., Tonta, M. A., Zoltkowski, B. Z., Meeker, W. F., Wendt, I. and Parkington, H. C. (2006). Pyeloureteric peristalsis: role of atypical smooth muscle cells and interstitial cells of Cajal-like cells as pacemakers. *J Physiol* 576, 695-705.

Lanner, F. and Rossant, J. (2010). The role of FGF/Erk signaling in pluripotent cells. *Development* 137, 3351-60.

Lawson, K. A., Meneses, J. J. and Pedersen, R. A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* 113, 891-911.

Le Douarin, N. M., Creuzet, S., Couly, G. and Dupin, E. (2004). Neural crest cell plasticity and its limits. *Development* 131, 4637-50.

Lee, E., Salic, A., Kruger, R., Heinrich, R. and Kirschner, M. W. (2003). The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol* 1, E10.

Lemaire, P., Bertrand, V. and Hudson, C. (2002). Early steps in the formation of neural tissue in ascidian embryos. *Dev Biol* 252, 151-69.

Lengerke, C. and Daley, G. Q. (2012). Caudal genes in blood development and leukemia. *Ann N Y Acad Sci* 1266, 47-54.

Lengerke, C., McKinney-Freeman, S., Naveiras, O., Yates, F., Wang, Y., Bansal, D. and Daley, G. Q. (2007). The cdx-hox pathway in hematopoietic stem cell formation from embryonic stem cells. *Ann N Y Acad Sci* 1106, 197-208.

Lengerke, C., Schmitt, S., Bowman, T. V., Jang, I. H., Maouche-Chretien, L., McKinney-Freeman, S., Davidson, A. J., Hammerschmidt, M., Rentzsch, F., Green, J. B. et al. (2008). BMP and Wnt specify hematopoietic fate by activation of the Cdx-Hox pathway. *Cell Stem Cell* 2, 72-82.

Levin, M. D., Lu, M. M., Petrenko, N. B., Hawkins, B. J., Gupta, T. H., Lang, D., Buckley, P. T., Jochems, J., Liu, F., Spurney, C. F. et al. (2009). Melanocyte-like cells in the heart and pulmonary veins contribute to atrial arrhythmia triggers. *J Clin Invest* 119, 3420-36.

Levine, A. J. and Brivanlou, A. H. (2007). Proposal of a model of mammalian neural induction. *Dev Biol* 308, 247-56.

Levy, C., Khaled, M. and Fisher, D. E. (2006). MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol Med* 12, 406-14.

Lewis, S. L., Khoo, P. L., De Young, R. A., Steiner, K., Wilcock, C., Mukhopadhyay, M., Westphal, H., Jamieson, R. V., Robb, L. and Tam, P. P. (2008). Dkk1 and Wnt3 interact to control head morphogenesis in the mouse. *Development* 135, 1791-801.

Li, B., Kuriyama, S., Moreno, M. and Mayor, R. (2009). The posteriorizing gene Gbx2 is a direct target of Wnt signalling and the earliest factor in neural crest induction. *Development* 136, 3267-78.

Li, J., Chen, F. and Epstein, J. A. (2000). Neural crest expression of Cre recombinase directed by the proximal Pax3 promoter in transgenic mice. *Genesis* 26, 162-4.

Li, J., Liu, K. C., Jin, F., Lu, M. M. and Epstein, J. A. (1999). Transgenic rescue of congenital heart disease and spina bifida in Splotch mice. *Development* 126, 2495-503.

Li, V. S., Ng, S. S., Boersema, P. J., Low, T. Y., Karthaus, W. R., Gerlach, J. P., Mohammed, S., Heck, A. J., Maurice, M. M., Mahmoudi, T. et al. (2012). Wnt signaling through inhibition of beta-catenin degradation in an intact Axin1 complex. *Cell* 149, 1245-56.

Lickert, H., Domon, C., Huls, G., Wehrle, C., Duluc, I., Clevers, H., Meyer, B. I., Freund, J. N. and Kemler, R. (2000). Wnt/(beta)-catenin signaling regulates the expression of the homeobox gene Cdx1 in embryonic intestine. *Development* 127, 3805-13.

Lickert, H. and Kemler, R. (2002). Functional analysis of cis-regulatory elements controlling initiation and maintenance of early Cdx1 gene expression in the mouse. *Dev Dyn* 225, 216-20.

Liem, K. F., Jr., Tremml, G. and Jessell, T. M. (1997). A role for the roof plate and its resident TGFbeta-related proteins in neuronal patterning in the dorsal spinal cord. *Cell* 91, 127-38.

Liem, K. F., Jr., Tremml, G., Roelink, H. and Jessell, T. M. (1995). Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm. *Cell* 82, 969-79.

Lin, W. M., Baker, A. C., Beroukhim, R., Winckler, W., Feng, W., Marmion, J. M., Laine, E., Greulich, H., Tseng, H., Gates, C. et al. (2008). Modeling genomic diversity and tumor dependency in malignant melanoma. *Cancer Res* 68, 664-73.

Lister, J. A., Cooper, C., Nguyen, K., Modrell, M., Grant, K. and Raible, D. W. (2006). Zebrafish Foxd3 is required for development of a subset of neural crest derivatives. *Dev Biol* 290, 92-104.

Litingtung, Y. and Chiang, C. (2000a). Control of Shh activity and signaling in the neural tube. *Dev Dyn* 219, 143-54.

Litingtung, Y. and Chiang, C. (2000b). Specification of ventral neuron types is mediated by an antagonistic interaction between Shh and Gli3. *Nat Neurosci* 3, 979-85.

Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G. H., Tan, Y., Zhang, Z., Lin, X. and He, X. (2002). Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell* 108, 837-47.

Liu, S., Liu, F., Schneider, A. E., St Amand, T., Epstein, J. A. and Gutstein, D. E. (2006a). Distinct cardiac malformations caused by absence of connexin 43 in the neural crest and in the non-crest neural tube. *Development* 133, 2063-73.

Liu, T., Branch, D. R. and Jin, T. (2006b). Pbx1 is a co-factor for Cdx-2 in regulating proglucagon gene expression in pancreatic A cells. *Mol Cell Endocrinol* 249, 140-9.

Liu, X., Luo, M., Xie, W., Wells, J. M., Goodheart, M. J. and Engelhardt, J. F. (2010). Sox17 modulates Wnt3A/beta-catenin-mediated transcriptional activation of the Lef-1 promoter. *Am J Physiol Lung Cell Mol Physiol* 299, L694-710.

Liu, Y., Helms, A. W. and Johnson, J. E. (2004). Distinct activities of Msx1 and Msx3 in dorsal neural tube development. *Development* 131, 1017-28.

Liu, Z., Merkurjev, D., Yang, F., Li, W., Oh, S., Friedman, M. J., Song, X., Zhang, F., Ma, Q., Ohgi, K. A. et al. (2014). Enhancer activation requires trans-recruitment of a mega transcription factor complex. *Cell* 159, 358-73.

Lohnes, D. (2003). The Cdx1 homeodomain protein: an integrator of posterior signaling in the mouse. *Bioessays* 25, 971-80.

Lorentz, O., Suh, E. R., Taylor, J. K., Boudreau, F. and Traber, P. G. (1999). CREB-binding [corrected] protein interacts with the homeodomain protein Cdx2 and enhances transcriptional activity. *J Biol Chem* 274, 7196-9.

Lowery, L. A. and Sive, H. (2004). Strategies of vertebrate neurulation and a re-evaluation of teleost neural tube formation. *Mech Dev* 121, 1189-97.

Lu, X., Borchers, A. G., Jolicoeur, C., Rayburn, H., Baker, J. C. and Tessier-Lavigne, M. (2004). PTK7/CCK-4 is a novel regulator of planar cell polarity in vertebrates. *Nature* 430, 93-8.

Lwigale, P. Y., Conrad, G. W. and Bronner-Fraser, M. (2004). Graded potential of neural crest to form cornea, sensory neurons and cartilage along the rostrocaudal axis. *Development* 131, 1979-91.

MacDonald, B. T., Tamai, K. and He, X. (2009). Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev Cell* 17, 9-26.

Macdonald, P. M. and Struhl, G. (1986). A molecular gradient in early Drosophila embryos and its role in specifying the body pattern. *Nature* 324, 537-45.

MacKenzie, A., Purdie, L., Davidson, D., Collinson, M. and Hill, R. E. (1997). Two enhancer domains control early aspects of the complex expression pattern of Msx1. *Mech Dev* 62, 29-40.

Maguire, L. H., Thomas, A. R. and Goldstein, A. M. (2014). Tumors of the neural crest: Common themes in development and cancer. *Dev Dyn*.

Mallo, G. V., Rechreche, H., Frigerio, J. M., Rocha, D., Zweibaum, A., Lacasa, M., Jordan, B. R., Dusetti, N. J., Dagorn, J. C. and Iovanna, J. L. (1997). Molecular cloning, sequencing and expression of the mRNA encoding human Cdx1 and Cdx2 homeobox. Down-regulation of Cdx1 and Cdx2 mRNA expression during colorectal carcinogenesis. *Int J Cancer* 74, 35-44.

Mansouri, A. and Gruss, P. (1998). Pax3 and Pax7 are expressed in commissural neurons and restrict ventral neuronal identity in the spinal cord. *Mech Dev* 78, 171-8.

Mansouri, A., Pla, P., Larue, L. and Gruss, P. (2001). Pax3 acts cell autonomously in the neural tube and somites by controlling cell surface properties. *Development* 128, 1995-2005.

Mansouri, A., Stoykova, A., Torres, M. and Gruss, P. (1996). Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice. *Development* 122, 831-8.

Marcelle, C., Stark, M. R. and Bronner-Fraser, M. (1997). Coordinate actions of BMPs, Wnts, Shh and noggin mediate patterning of the dorsal somite. *Development* 124, 3955-63.

Maretto, S., Cordenonsi, M., Dupont, S., Braghetta, P., Broccoli, V., Hassan, A. B., Volpin, D., Bressan, G. M. and Piccolo, S. (2003). Mapping Wnt/beta-catenin signaling during mouse development and in colorectal tumors. *Proc Natl Acad Sci U S A* 100, 3299-304.

Margalit, Y., Yarus, S., Shapira, E., Gruenbaum, Y. and Fainsod, A. (1993). Isolation and characterization of target sequences of the chicken CdxA homeobox gene. *Nucleic Acids Res* 21, 4915-22.

Mari, L., Milano, F., Parikh, K., Straub, D., Everts, V., Hoeben, K. K., Fockens, P., Buttar, N. S. and Krishnadath, K. K. (2014). A pSMAD/CDX2 complex is essential for the intestinalization of epithelial metaplasia. *Cell Rep* 7, 1197-210.

Marikawa, Y., Tamashiro, D. A., Fujita, T. C. and Alarcon, V. B. (2009). Aggregated P19 mouse embryonal carcinoma cells as a simple in vitro model to study the molecular regulations of mesoderm formation and axial elongation morphogenesis. *Genesis* 47, 93-106.

Marom, K., Shapira, E. and Fainsod, A. (1997). The chicken caudal genes establish an anterior-posterior gradient by partially overlapping temporal and spatial patterns of expression. *Mech Dev* 64, 41-52.

Marson, A., Levine, S. S., Cole, M. F., Frampton, G. M., Brambrink, T., Johnstone, S., Guenther, M. G., Johnston, W. K., Wernig, M., Newman, J. et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134, 521-33.

Mas, C. and Ruiz i Altaba, A. (2010). Small molecule modulation of HH-GLI signaling: current leads, trials and tribulations. *Biochem Pharmacol* 80, 712-23.

Mathieu, M., Bourges, E., Caron, F. and Piussan, C. (1990). [Waardenburg's syndrome and severe cyanotic cardiopathy]. *Arch Fr Pediatr* 47, 657-9.

Mathis, J. M., Simmons, D. M., He, X., Swanson, L. W. and Rosenfeld, M. G. (1992). Brain 4: a novel mammalian POU domain transcription factor exhibiting restricted brain-specific expression. *EMBO J* 11, 2551-61.

Matise, M. P., Epstein, D. J., Park, H. L., Platt, K. A. and Joyner, A. L. (1998). Gli2 is required for induction of floor plate and adjacent cells, but not most ventral neurons in the mouse central nervous system. *Development* 125, 2759-70.

McGrew, L. L., Lai, C. J. and Moon, R. T. (1995). Specification of the anteroposterior neural axis through synergistic interaction of the Wnt signaling cascade with noggin and follistatin. *Dev Biol* 172, 337-42.

McMahon, J. A., Takada, S., Zimmerman, L. B., Fan, C. M., Harland, R. M. and McMahon, A. P. (1998). Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes Dev* 12, 1438-52.

Medic, S. and Ziman, M. (2009). PAX3 across the spectrum: from melanoblast to melanoma. *Crit Rev Biochem Mol Biol* 44, 85-97.

Mehra-Chaudhary, R., Matsui, H. and Raghow, R. (2001). Msx3 protein recruits histone deacetylase to down-regulate the Msx1 promoter. *Biochem J* 353, 13-22.

Merello, E., Mascelli, S., Raso, A., Piatelli, G., Consales, A., Cama, A., Kibar, Z., Capra, V. and Marco, P. D. (2014). Expanding the mutational spectrum associated to neural tube defects: Literature revision and description of novel VANGL1 mutations. *Birth Defects Res A Clin Mol Teratol.*

Merzdorf, C. S. (2007). Emerging roles for zic genes in early development. *Dev Dyn* 236, 922-40.

Meulemans, D. and Bronner-Fraser, M. (2004). Gene-regulatory interactions in neural crest evolution and development. *Dev Cell* 7, 291-9.

Meyer, B. I. and Gruss, P. (1993). Mouse Cdx-1 expression during gastrulation. *Development* 117, 191-203.

Milet, C., Maczkowiak, F., Roche, D. D. and Monsoro-Burq, A. H. (2013). Pax3 and Zic1 drive induction and differentiation of multipotent, migratory, and functional neural crest in Xenopus embryos. *Proc Natl Acad Sci U S A* 110, 5528-33.

Milewski, R. C., Chi, N. C., Li, J., Brown, C., Lu, M. M. and Epstein, J. A. (2004). Identification of minimal enhancer elements sufficient for Pax3 expression in neural crest and implication of Tead2 as a regulator of Pax3. *Development* 131, 829-37.

Miller, K. A., Barrow, J., Collinson, J. M., Davidson, S., Lear, M., Hill, R. E. and Mackenzie, A. (2007). A highly conserved Wnt-dependent TCF4 binding site within the proximal enhancer of the anti-myogenic Msx1 gene supports expression within Pax3-expressing limb bud muscle precursor cells. *Dev Biol* 311, 665-78.

Miller, K. A., Davidson, S., Liaros, A., Barrow, J., Lear, M., Heine, D., Hoppler, S. and MacKenzie, A. (2008). Prediction and characterisation of a highly conserved, remote and cAMP responsive enhancer that regulates Msx1 gene expression in cardiac neural crest and outflow tract. *Dev Biol* 317, 686-94.

Mita, K. and Fujiwara, S. (2007). Nodal regulates neural tube formation in the Ciona intestinalis embryo. *Dev Genes Evol* 217, 593-601.

Mitchell, L. E. (2005). Epidemiology of neural tube defects. *Am J Med Genet C Semin Med Genet* 135C, 88-94.

Mitchell, P. J., Timmons, P. M., Hebert, J. M., Rigby, P. W. and Tjian, R. (1991). Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. *Genes Dev* 5, 105-19.

Mitchelmore, C., Troelsen, J. T., Spodsberg, N., Sjostrom, H. and Noren, O. (2000). Interaction between the homeodomain proteins Cdx2 and HNF1alpha mediates expression of the lactase-phlorizin hydrolase gene. *Biochem J* 346 Pt 2, 529-35.

Miyazono, K., Kamiya, Y. and Morikawa, M. (2010). Bone morphogenetic protein receptors and signal transduction. *J Biochem* 147, 35-51.

Mizugishi, K., Aruga, J., Nakata, K. and Mikoshiba, K. (2001). Molecular properties of Zic proteins as transcriptional regulators and their relationship to GLI proteins. *J Biol Chem* 276, 2180-8.

Mizugishi, K., Hatayama, M., Tohmonda, T., Ogawa, M., Inoue, T., Mikoshiba, K. and Aruga, J. (2004). Myogenic repressor I-mfa interferes with the function of Zic family proteins. *Biochem Biophys Res Commun* 320, 233-40.

Mlodzik, M., Fjose, A. and Gehring, W. J. (1985). Isolation of caudal, a Drosophila homeo box-containing gene with maternal expression, whose transcripts form a concentration gradient at the pre-blastoderm stage. *EMBO J* 4, 2961-9.

Mlodzik, M. and Gehring, W. J. (1987). Expression of the caudal gene in the germ line of Drosophila: formation of an RNA and protein gradient during early embryogenesis. *Cell* 48, 465-78.

Molotkova, N., Molotkov, A., Sirbu, I. O. and Duester, G. (2005). Requirement of mesodermal retinoic acid generated by Raldh2 for posterior neural transformation. *Mech Dev* 122, 145-55.

Monsoro-Burq, A. H., Fletcher, R. B. and Harland, R. M. (2003). Neural crest induction by paraxial mesoderm in Xenopus embryos requires FGF signals. *Development* 130, 3111-24.

Monsoro-Burq, A. H., Wang, E. and Harland, R. (2005). Msx1 and Pax3 cooperate to mediate FGF8 and WNT signals during Xenopus neural crest induction. *Dev Cell* 8, 167-78.

Montero-Balaguer, M., Lang, M. R., Sachdev, S. W., Knappmeyer, C., Stewart, R. A., De La Guardia, A., Hatzopoulos, A. K. and Knapik, E. W. (2006). The mother superior mutation ablates foxd3 activity in neural crest progenitor cells and depletes neural crest derivatives in zebrafish. *Dev Dyn* 235, 3199-212.

Monuki, E. S., Kuhn, R., Weinmaster, G., Trapp, B. D. and Lemke, G. (1990). Expression and activity of the POU transcription factor SCIP. *Science* 249, 1300-3.

Moore, S., Ribes, V., Terriente, J., Wilkinson, D., Relaix, F. and Briscoe, J. (2013). Distinct regulatory mechanisms act to establish and maintain Pax3 expression in the developing neural tube. *PLoS Genet* 9, e1003811.

Moreno, E. and Morata, G. (1999). Caudal is the Hox gene that specifies the most posterior Drosophile segment. *Nature* 400, 873-7.

Morris, G. L., O'Shea K.S. (1983). Anomalies of neuroepithelial cell associations in the splotch mutant embryo. *Dev Brain Res* 9, 408-410.

Moustakas, A. and Heldin, C. H. (2009). The regulation of TGFbeta signal transduction. *Development* 136, 3699-714.

Mukhopadhyay, M., Shtrom, S., Rodriguez-Esteban, C., Chen, L., Tsukui, T., Gomer, L., Dorward, D. W., Glinka, A., Grinberg, A., Huang, S. P. et al. (2001). Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse. *Dev Cell* 1, 423-34.

Mundell, N. A. and Labosky, P. A. (2011). Neural crest stem cell multipotency requires Foxd3 to maintain neural potential and repress mesenchymal fates. *Development* 138, 641-52.

Munoz-Sanjuan, I. and Brivanlou, A. H. (2002). Neural induction, the default model and embryonic stem cells. *Nat Rev Neurosci* 3, 271-80.

Nagai, T., Aruga, J., Minowa, O., Sugimoto, T., Ohno, Y., Noda, T. and Mikoshiba, K. (2000). Zic2 regulates the kinetics of neurulation. *Proc Natl Acad Sci U S A* 97, 1618-23.

Nagai, T., Aruga, J., Takada, S., Gunther, T., Sporle, R., Schughart, K. and Mikoshiba, K. (1997). The expression of the mouse Zic1, Zic2, and Zic3 gene suggests an essential role for Zic genes in body pattern formation. *Dev Biol* 182, 299-313.

Nagy, A., Gertsenstein, M., Vintersten, K. and Behringer, R. (2003). Manipulating the mouse embryo, (ed. C. S. H. laboratory). NY: Cold Spring Harbor.

Nakao, A., Afrakhte, M., Moren, A., Nakayama, T., Christian, J. L., Heuchel, R., Itoh, S., Kawabata, M., Heldin, N. E., Heldin, C. H. et al. (1997). Identification of Smad7, a TGFbeta-inducible antagonist of TGF-beta signalling. *Nature* 389, 631-5.

Natoli, T. A., Ellsworth, M. K., Wu, C., Gross, K. W. and Pruitt, S. C. (1997). Positive and negative DNA sequence elements are required to establish the pattern of Pax3 expression. *Development* 124, 617-26.

Nelms, B. L., Pfaltzgraff, E. R. and Labosky, P. A. (2011). Functional interaction between Foxd3 and Pax3 in cardiac neural crest development. *Genesis* 49, 10-23.

Niehrs, C. (1999). Head in the WNT: the molecular nature of Spemann's head organizer. *Trends Genet* 15, 314-9.

Nieuwkoop, P. (1954). Neural activation and transformation in explants of competent ectoderm under the influence of fragments of anterior notochord in urodeles. *J Embryol Exp Morphol* 2, 175-193.

Nikitina, N., Sauka-Spengler, T. and Bronner-Fraser, M. (2008). Dissecting early regulatory relationships in the lamprey neural crest gene network. *Proc Natl Acad Sci U S A* 105, 20083-8.

Nikitina, N. V. and Bronner-Fraser, M. (2009). Gene regulatory networks that control the specification of neural-crest cells in the lamprey. *Biochim Biophys Acta* 1789, 274-8.

Nishida, H. and Satoh, N. (1989). Determination and regulation in the pigment cell lineage of the ascidian embryo. *Dev Biol* 132, 355-67.

Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyama, K., Utsunomiya, J., Baba, S. and Hedge, P. (1991). Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* 253, 665-9.

Nitzan, E., Krispin, S., Pfaltzgraff, E. R., Klar, A., Labosky, P. A. and Kalcheim, C. (2013). A dynamic code of dorsal neural tube genes regulates the segregation between neurogenic and melanogenic neural crest cells. *Development* 140, 2269-79.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R. and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* 123, 917-29.

Nolan, P. M., Peters, J., Strivens, M., Rogers, D., Hagan, J., Spurr, N., Gray, I. C., Vizor, L., Brooker, D., Whitehill, E. et al. (2000). A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat Genet* 25, 440-3.

Nordeen, S. K. (1988). Luciferase reporter gene vectors for analysis of promoter and enhancers. *Biotechniques* 6, 454-458.

Nordstrom, U., Maier, E., Jessell, T. M. and Edlund, T. (2006). An early role for WNT signaling in specifying neural patterns of Cdx and Hox gene expression and motor neuron subtype identity. *PLoS Biol* 4, e252.

Nusse, R. and Varmus, H. E. (1982). Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* 31, 99-109.

Nusslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature* 287, 795-801.

Nyabi, O., Naessens, M., Haigh, K., Gembarska, A., Goossens, S., Maetens, M., De Clercq, S., Drogat, B., Haenebalcke, L., Bartunkova, S. et al. (2009). Efficient mouse transgenesis using Gateway-compatible ROSA26 locus targeting vectors and F1 hybrid ES cells. *Nucleic Acids Res* 37, e55.

Nye, J. S., McLone, D. G., Charrow, J. and Hayes, E. A. (1999). Neural crest anomaly syndromes in children with spina bifida. *Teratology* 60, 179-89.

Nyholm, M. K., Abdelilah-Seyfried, S. and Grinblat, Y. (2009). A novel genetic mechanism regulates dorsolateral hinge-point formation during zebrafish cranial neurulation. *J Cell Sci* 122, 2137-48.

Nyholm, M. K., Wu, S. F., Dorsky, R. I. and Grinblat, Y. (2007). The zebrafish zic2a-zic5 gene pair acts downstream of canonical Wnt signaling to control cell proliferation in the developing tectum. *Development* 134, 735-46.

Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. and Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 87, 953-9.

Ohneda, K., Mirmira, R. G., Wang, J., Johnson, J. D. and German, M. S. (2000). The homeodomain of PDX-1 mediates multiple protein-protein interactions in the formation of a transcriptional activation complex on the insulin promoter. *Mol Cell Biol* 20, 900-11.

Olaopa, M., Zhou, H. M., Snider, P., Wang, J., Schwartz, R. J., Moon, A. M. and Conway, S. J. (2011). Pax3 is essential for normal cardiac neural crest morphogenesis but is not required during migration nor outflow tract septation. *Dev Biol* 356, 308-22.

Opdecamp, K., Nakayama, A., Nguyen, M. T., Hodgkinson, C. A., Pavan, W. J. and Arnheiter, H. (1997). Melanocyte development in vivo and in neural crest cell cultures: crucial dependence on the Mitf basic-helix-loop-helix-zipper transcription factor. *Development* 124, 2377-86.

Osawa, M., Egawa, G., Mak, S. S., Moriyama, M., Freter, R., Yonetani, S., Beermann, F. and Nishikawa, S. (2005). Molecular characterization of melanocyte stem cells in their niche. *Development* 132, 5589-99.

Ozair, M. Z., Kintner, C. and Brivanlou, A. H. (2013). Neural induction and early patterning in vertebrates. *Wiley Interdiscip Rev Dev Biol* 2, 479-98.

Pan, G., Li, J., Zhou, Y., Zheng, H. and Pei, D. (2006). A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB J* 20, 1730-2.

Pani, L., Horal, M. and Loeken, M. R. (2002). Rescue of neural tube defects in Pax-3-deficient embryos by p53 loss of function: implications for Pax-3- dependent development and tumorigenesis. *Genes Dev* 16, 676-80.

Papalopulu, N. and Kintner, C. (1996). A posteriorising factor, retinoic acid, reveals that anteroposterior patterning controls the timing of neuronal differentiation in Xenopus neuroectoderm. *Development* 122, 3409-18.

Park, H. L., Bai, C., Platt, K. A., Matise, M. P., Beeghly, A., Hui, C. C., Nakashima, M. and Joyner, A. L. (2000). Mouse Gli1 mutants are viable but have defects in SHH signaling in combination with a Gli2 mutation. *Development* 127, 1593-605.

Partanen, J., Schwartz, L. and Rossant, J. (1998). Opposite phenotypes of hypomorphic and Y766 phosphorylation site mutations reveal a function for Fgfr1 in anteroposterior patterning of mouse embryos. *Genes Dev* 12, 2332-44.

Passeron, T., Valencia, J. C., Namiki, T., Vieira, W. D., Passeron, H., Miyamura, Y. and Hearing, V. J. (2009). Upregulation of SOX9 inhibits the growth of human and mouse melanomas and restores their sensitivity to retinoic acid. *J Clin Invest* 119, 954-63.

Pavan, W. J. and Raible, D. W. (2012). Specification of neural crest into sensory neuron and melanocyte lineages. *Dev Biol* 366, 55-63.

Pelegri, F. and Maischein, H. M. (1998). Function of zebrafish beta-catenin and TCF-3 in dorsoventral patterning. *Mech Dev* 77, 63-74.

Perantoni, A. O., Timofeeva, O., Naillat, F., Richman, C., Pajni-Underwood, S., Wilson, C., Vainio, S., Dove, L. F. and Lewandoski, M. (2005). Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. *Development* 132, 3859-71.

Perea-Gomez, A., Lawson, K. A., Rhinn, M., Zakin, L., Brulet, P., Mazan, S. and Ang, S. L. (2001). Otx2 is required for visceral endoderm movement and for the restriction of posterior signals in the epiblast of the mouse embryo. *Development* 128, 753-65.

Perea-Gomez, A., Shawlot, W., Sasaki, H., Behringer, R. R. and Ang, S. (1999). HNF3beta and Lim1 interact in the visceral endoderm to regulate primitive streak formation and anterior-posterior polarity in the mouse embryo. *Development* 126, 4499-511.

Perea-Gomez, A., Vella, F. D., Shawlot, W., Oulad-Abdelghani, M., Chazaud, C., Meno, C., Pfister, V., Chen, L., Robertson, E., Hamada, H. et al. (2002). Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. *Dev Cell* 3, 745-56.

Petersen, C. P. and Reddien, P. W. (2009). Wnt signaling and the polarity of the primary body axis. *Cell* 139, 1056-68.

Peterson, K. A., Nishi, Y., Ma, W., Vedenko, A., Shokri, L., Zhang, X., McFarlane, M., Baizabal, J. M., Junker, J. P., van Oudenaarden, A. et al. (2012). Neural-specific Sox2 input and differential Gli-binding affinity provide context and positional information in Shh-directed neural patterning. *Genes Dev* 26, 2802-16.

Petropoulos, H. and Skerjanc, I. S. (2002). Beta-catenin is essential and sufficient for skeletal myogenesis in P19 cells. *J Biol Chem* 277, 15393-9.

Phelan, S. A., Ito, M. and Loeken, M. R. (1997). Neural tube defects in embryos of diabetic mice: role of the Pax-3 gene and apoptosis. *Diabetes* 46, 1189-97.

Piccolo, S., Sasai, Y., Lu, B. and De Robertis, E. M. (1996). Dorsoventral patterning in Xenopus: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* 86, 589-98.

Pilon, N., Oh, K., Sylvestre, J. R., Bouchard, N., Savory, J. and Lohnes, D. (2006). Cdx4 is a direct target of the canonical Wnt pathway. *Dev Biol* 289, 55-63.

Pilon, N., Oh, K., Sylvestre, J. R., Savory, J. G. and Lohnes, D. (2007). Wnt signaling is a key mediator of Cdx1 expression in vivo. *Development* 134, 2315-23.

Pilon, N., Raiwet, D., Viger, R. S. and Silversides, D. W. (2008). Novel pre- and post-gastrulation expression of Gata4 within cells of the inner cell mass and migratory neural crest cells. *Dev Dyn* 237, 1133-43.

Pilozzi, E., Onelli, M. R., Ziparo, V., Mercantini, P. and Ruco, L. (2004). CDX1 expression is reduced in colorectal carcinoma and is associated with promoter hypermethylation. *J Pathol* 204, 289-95.

Pollard, S. L. and Holland, P. W. (2000). Evidence for 14 homeobox gene clusters in human genome ancestry. *Curr Biol* 10, 1059-62.

Popperl, H., Schmidt, C., Wilson, V., Hume, C. R., Dodd, J., Krumlauf, R. and Beddington, R. S. (1997). Misexpression of Cwnt8C in the mouse induces an ectopic embryonic axis and causes a truncation of the anterior neuroectoderm. *Development* 124, 2997-3005.

Pourebrahim, R., Houtmeyers, R., Ghogomu, S., Janssens, S., Thelie, A., Tran, H. T., Langenberg, T., Vleminckx, K., Bellefroid, E., Cassiman, J. J. et al. (2011). Transcription factor Zic2 inhibits Wnt/beta-catenin protein signaling. *J Biol Chem* 286, 37732-40.

Pourquie, O. (2003). The segmentation clock: converting embryonic time into spatial pattern. *Science* 301, 328-30.

Prinos, P., Joseph, S., Oh, K., Meyer, B. I., Gruss, P. and Lohnes, D. (2001). Multiple pathways governing Cdx1 expression during murine development. *Dev Biol* 239, 257-69.

Pruitt, S. C., Bussman, A., Maslov, A. Y., Natoli, T. A. and Heinaman, R. (2004). Hox/Pbx and Brn binding sites mediate Pax3 expression in vitro and in vivo. *Gene Expr Patterns* 4, 671-85.

Qin, Y., Kong, L. K., Poirier, C., Truong, C., Overbeek, P. A. and Bishop, C. E. (2004). Long-range activation of Sox9 in Odd Sex (Ods) mice. *Hum Mol Genet* 13, 1213-8.

Raible, D. W., Wood, A., Hodsdon, W., Henion, P. D., Weston, J. A. and Eisen, J. S. (1992). Segregation and early dispersal of neural crest cells in the embryonic zebrafish. *Dev Dyn* 195, 29-42.

Rave-Harel, N., Miller, N. L., Givens, M. L. and Mellon, P. L. (2005). The Groucho-related gene family regulates the gonadotropin-releasing hormone gene through interaction with the homeodomain proteins MSX1 and OCT1. *J Biol Chem* 280, 30975-83.

Read, A. P. and Newton, V. E. (1997). Waardenburg syndrome. *J Med Genet* 34, 656-65.

Reddi, A. H. and Huggins, C. (1972). Biochemical sequences in the transformation of normal fibroblasts in adolescent rats. *Proc Natl Acad Sci U S A* 69, 1601-5.

Reece-Hoyes, J. S., Keenan, I. D. and Isaacs, H. V. (2002). Cloning and expression of the Cdx family from the frog Xenopus tropicalis. *Dev Dyn* 223, 134-40.

Relaix, F., Rocancourt, D., Mansouri, A. and Buckingham, M. (2004). Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev* 18, 1088-105.

Riddle, R. D., Johnson, R. L., Laufer, E. and Tabin, C. (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* 75, 1401-16.

Rijsewijk, F., Schuermann, M., Wagenaar, E., Parren, P., Weigel, D. and Nusse, R. (1987). The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell* 50, 649-57.

Rings, E. H., Boudreau, F., Taylor, J. K., Moffett, J., Suh, E. R. and Traber, P. G. (2001). Phosphorylation of the serine 60 residue within the Cdx2 activation domain mediates its transactivation capacity. *Gastroenterology* 121, 1437-50.

Rohatgi, R., Milenkovic, L. and Scott, M. P. (2007). Patched1 regulates hedgehog signaling at the primary cilium. *Science* 317, 372-6.

Rohr, K. B., Schulte-Merker, S. and Tautz, D. (1999). Zebrafish zic1 expression in brain and somites is affected by BMP and hedgehog signalling. *Mech Dev* 85, 147-59.

Roose, J., Molenaar, M., Peterson, J., Hurenkamp, J., Brantjes, H., Moerer, P., van de Wetering, M., Destree, O. and Clevers, H. (1998). The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature* 395, 608-12.

Rossant, J. and Tam, P. P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* 136, 701-13.

Rossant, J., Zirngibl, R., Cado, D., Shago, M. and Giguere, V. (1991). Expression of a retinoic acid response element-hsplacZ transgene defines specific domains of transcriptional activity during mouse embryogenesis. *Genes Dev* 5, 1333-44.

Russell, W. L. (1947). Splotch, a new mutation in the house mouse *Mus musculus*. *Genetics* 32, 107.

Saegusa, M., Hashimura, M., Kuwata, T., Hamano, M., Wani, Y. and Okayasu, I. (2007). A functional role of Cdx2 in beta-catenin signaling during transdifferentiation in endometrial carcinomas. *Carcinogenesis* 28, 1885-92.

Sakai, Y., Meno, C., Fujii, H., Nishino, J., Shiratori, H., Saijoh, Y., Rossant, J. and Hamada, H. (2001). The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the anterio-posterior axis within the mouse embryo. *Genes Dev* 15, 213-25.

Salari, K., Spulak, M. E., Cuff, J., Forster, A. D., Giacomini, C. P., Huang, S., Ko, M. E., Lin, A. Y., van de Rijn, M. and Pollack, J. R. (2012). CDX2 is an amplified lineage-survival oncogene in colorectal cancer. *Proc Natl Acad Sci U S A* 109, E3196-205.

Salero, E., Perez-Sen, R., Aruga, J., Gimenez, C. and Zafra, F. (2001). Transcription factors Zic1 and Zic2 bind and transactivate the apolipoprotein E gene promoter. *J Biol Chem* 276, 1881-8.

Sanchez-Ferras, O., Bernas, G., Laberge-Perrault, E. and Pilon, N. (2014). Induction and dorsal restriction of Paired-box 3 (Pax3) gene expression in the caudal neuroectoderm is mediated by integration of multiple pathways on a short neural crest enhancer. *Biochim Biophys Acta* 1839, 546-58.

Sanchez-Ferras, O., Coutaud, B., Djavanbakht Samani, T., Tremblay, I., Souchkova, O. and Pilon, N. (2012). Caudal-related homeobox (Cdx) protein-dependent integration of canonical Wnt signaling on paired-box 3 (Pax3) neural crest enhancer. *J Biol Chem* 287, 16623-35.

Sasai, N., Mizuseki, K. and Sasai, Y. (2001). Requirement of FoxD3-class signaling for neural crest determination in Xenopus. *Development* 128, 2525-36.

Sasai, Y., Lu, B., Steinbeisser, H., Geissert, D., Gont, L. K. and De Robertis, E. M. (1994). Xenopus chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell* 79, 779-90.

Sato, T., Sasai, N. and Sasai, Y. (2005). Neural crest determination by co-activation of Pax3 and Zic1 genes in Xenopus ectoderm. *Development* 132, 2355-63.

Satokata, I., Ma, L., Ohshima, H., Bei, M., Woo, I., Nishizawa, K., Maeda, T., Takano, Y., Uchiyama, M., Heaney, S. et al. (2000). Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. *Nat Genet* 24, 391-5.

Satokata, I. and Maas, R. (1994). Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nat Genet* 6, 348-56.

Sauka-Spengler, T. and Bronner-Fraser, M. (2008). A gene regulatory network orchestrates neural crest formation. *Nat Rev Mol Cell Biol* 9, 557-68.

Sauka-Spengler, T., Meulemans, D., Jones, M. and Bronner-Fraser, M. (2007). Ancient evolutionary origin of the neural crest gene regulatory network. *Dev Cell* 13, 405-20.

Savory, J. G., Bouchard, N., Pierre, V., Rijli, F. M., De Repentigny, Y., Kothary, R. and Lohnes, D. (2009a). Cdx2 regulation of posterior development through non-Hox targets. *Development* 136, 4099-110.

Savory, J. G., Mansfield, M., Rijli, F. M. and Lohnes, D. (2011a). Cdx mediates neural tube closure through transcriptional regulation of the planar cell polarity gene Ptk7. *Development* 138, 1361-70.

Savory, J. G., Mansfield, M., St Louis, C. and Lohnes, D. (2011b). Cdx4 is a Cdx2 target gene. *Mech Dev* 128, 41-8.

Savory, J. G., Pilon, N., Grainger, S., Sylvestre, J. R., Beland, M., Houle, M., Oh, K. and Lohnes, D. (2009b). Cdx1 and Cdx2 are functionally equivalent in vertebral patterning. *Dev Biol* 330, 114-22.

Sawada, A., Nishizaki, Y., Sato, H., Yada, Y., Nakayama, R., Yamamoto, S., Nishioka, N., Kondoh, H. and Sasaki, H. (2005). Tead proteins activate the Foxa2 enhancer in the node in cooperation with a second factor. *Development* 132, 4719-29.

Schoenwolf, G. C. (1984). Histological and ultrastructural studies of secondary neurulation in mouse embryos. *Am J Anat* 169, 361-76.

Schorle, H., Meier, P., Buchert, M., Jaenisch, R. and Mitchell, P. J. (1996). Transcription factor AP-2 essential for cranial closure and craniofacial development. *Nature* 381, 235-8.

Schwarz-Romond, T., Merrifield, C., Nichols, B. J. and Bienz, M. (2005). The Wnt signalling effector Dishevelled forms dynamic protein assemblies rather than stable associations with cytoplasmic vesicles. *J Cell Sci* 118, 5269-77.

Scott, M. P., Tamkun, J. W. and Hartzell, G. W., 3rd. (1989). The structure and function of the homeodomain. *Biochim Biophys Acta* 989, 25-48.

Seifert, J. R. and Mlodzik, M. (2007). Frizzled/PCP signalling: a conserved mechanism regulating cell polarity and directed motility. *Nat Rev Genet* 8, 126-38.

Serbedzija, G. N., Bronner-Fraser, M. and Fraser, S. E. (1989). A vital dye analysis of the timing and pathways of avian trunk neural crest cell migration. *Development* 106, 809-16.

Serbedzija, G. N., Fraser, S. E. and Bronner-Fraser, M. (1990). Pathways of trunk neural crest cell migration in the mouse embryo as revealed by vital dye labelling. *Development* 108, 605-12.

Serbedzija, G. N. and McMahon, A. P. (1997). Analysis of neural crest cell migration in Splotch mice using a neural crest-specific LacZ reporter. *Dev Biol* 185, 139-47.

Shakhova, O., Zingg, D., Schaefer, S. M., Hari, L., Civenni, G., Blunschi, J., Claudinot, S., Okoniewski, M., Beermann, F., Mihic-Probst, D. et al. (2012). Sox10 promotes the formation and maintenance of giant congenital naevi and melanoma. *Nat Cell Biol* 14, 882-90.

Shetty, S., Takahashi, T., Matsui, H., Ayengar, R. and Raghow, R. (1999). Transcriptional autorepression of Msx1 gene is mediated by interactions of Msx1 protein with a multi-protein transcriptional complex containing TATA-binding protein, Sp1 and cAMP-response-element-binding protein-binding protein (CBP/p300). *Biochem J* 339 (Pt 3), 751-8.

Shi, F., Cheng, Y. F., Wang, X. L. and Edge, A. S. (2010). Beta-catenin up-regulates Atoh1 expression in neural progenitor cells by interaction with an Atoh1 3' enhancer. *J Biol Chem* 285, 392-400.

Shibamoto, S., Higano, K., Takada, R., Ito, F., Takeichi, M. and Takada, S. (1998). Cytoskeletal reorganization by soluble Wnt-3a protein signalling. *Genes Cells* 3, 659-70.

Shimizu, T., Bae, Y. K. and Hibi, M. (2006). Cdx-Hox code controls competence for responding to Fgfs and retinoic acid in zebrafish neural tissue. *Development* 133, 4709-19.

Shimizu, T., Bae, Y. K., Muraoka, O. and Hibi, M. (2005). Interaction of Wnt and caudal-related genes in zebrafish posterior body formation. *Dev Biol* 279, 125-41.

Shir-Shapira, H., Sharabany, J., Filderman, M., Ideses, D., Ovadia-Shochat, A., Mannervik, M. and Juven-Gershon, T. (2015). Structure-Function Analysis of the Drosophila melanogaster Caudal Transcription Factor Provides Insights into Core Promoter-preferential Activation. *J Biol Chem* 290, 17293-305.

Shoguchi, E., Hamaguchi, M. and Satoh, N. (2008). Genome-wide network of regulatory genes for construction of a chordate embryo. *Dev Biol* 316, 498-509.

Simoes-Costa, M. S., McKeown, S. J., Tan-Cabugao, J., Sauka-Spengler, T. and Bronner, M. E. (2012). Dynamic and differential regulation of stem cell factor FoxD3 in the neural crest is Encrypted in the genome. *PLoS Genet* 8, e1003142.

Simons, M. and Mlodzik, M. (2008). Planar cell polarity signaling: from fly development to human disease. *Annu Rev Genet* 42, 517-40.

Sinnberg, T., Menzel, M., Ewerth, D., Sauer, B., Schwarz, M., Schaller, M., Garbe, C. and Schittek, B. (2011). beta-Catenin signaling increases during melanoma progression and promotes tumor cell survival and chemoresistance. *PLoS One* 6, e23429.

Skromne, I., Thorsen, D., Hale, M., Prince, V. E. and Ho, R. K. (2007). Repression of the hindbrain developmental program by Cdx factors is required for the specification of the vertebrate spinal cord. *Development* 134, 2147-58.

Smith, W. C. and Harland, R. M. (1992). Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in Xenopus embryos. *Cell* 70, 829-40.

Smukler, S. R., Runciman, S. B., Xu, S. and van der Kooy, D. (2006). Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. *J Cell Biol* 172, 79-90.

Solberg, N., Machon, O., Machonova, O. and Krauss, S. (2012). Mouse Tcf3 represses canonical Wnt signaling by either competing for beta-catenin binding or through occupation of DNA-binding sites. *Mol Cell Biochem* 365, 53-63.

Spemann, H. a. M., H. (1924). Uber Induction von Embryonanlagen durch Implantation Artfremder Organis Atoren. Wilhelm Roux. *Arch Entwicklungsmech Org* 100, 599-638.

Srinivas, S., Watanabe, T., Lin, C. S., William, C. M., Tanabe, Y., Jessell, T. M. and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* 1, 4.

Steel, K. P., Davidson, D. R. and Jackson, I. J. (1992). TRP-2/DT, a new early melanoblast marker, shows that steel growth factor (c-kit ligand) is a survival factor. *Development* 115, 1111-9.

Stern, C. D. (2005). Neural induction: old problem, new findings, yet more questions. *Development* 132, 2007-21.

Strumpf, D., Mao, C. A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F. and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* 132, 2093-102.

Stuhlmiller, T. J. and Garcia-Castro, M. I. (2012). Current perspectives of the signaling pathways directing neural crest induction. *Cell Mol Life Sci*.

Sturgeon, K., Kaneko, T., Biemann, M., Gauthier, A., Chawengsaksophak, K. and Cordes, S. P. (2011). Cdx1 refines positional identity of the vertebrate hindbrain by directly repressing Mafb expression. *Development* 138, 65-74.

Subramanian, V., Meyer, B. I. and Gruss, P. (1995). Disruption of the murine homeobox gene Cdx1 affects axial skeletal identities by altering the mesodermal expression domains of Hox genes. *Cell* 83, 641-53.

Sugitani, Y., Nakai, S., Minowa, O., Nishi, M., Jishage, K., Kawano, H., Mori, K., Ogawa, M. and Noda, T. (2002). Brn-1 and Brn-2 share crucial roles in the production and positioning of mouse neocortical neurons. *Genes Dev* 16, 1760-5.

Suh, E., Chen, L., Taylor, J. and Traber, P. G. (1994). A homeodomain protein related to caudal regulates intestine-specific gene transcription. *Mol Cell Biol* 14, 7340-51.

Sun, X., Meyers, E. N., Lewandoski, M. and Martin, G. R. (1999). Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. *Genes Dev* 13, 1834-46.

Sutton, J., Costa, R., Klug, M., Field, L., Xu, D., Largaespada, D. A., Fletcher, C. F., Jenkins, N. A., Copeland, N. G., Klemsz, M. et al. (1996). Genesis, a winged helix transcriptional repressor with expression restricted to embryonic stem cells. *J Biol Chem* 271, 23126-33.

Suzuki, A., Ueno, N. and Hemmati-Brivanlou, A. (1997). Xenopus msx1 mediates epidermal induction and neural inhibition by BMP4. *Development* 124, 3037-44.

Tachibana, M. (1999). Sound needs sound melanocytes to be heard. *Pigment Cell Res* 12, 344-54.

Takada, R., Satomi, Y., Kurata, T., Ueno, N., Norioka, S., Kondoh, H., Takao, T. and Takada, S. (2006). Monounsaturated fatty acid modification of Wnt protein: its role in Wnt secretion. *Dev Cell* 11, 791-801.

Takada, S., Stark, K. L., Shea, M. J., Vassileva, G., McMahon, J. A. and McMahon, A. P. (1994). Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes Dev* 8, 174-89.

Takahashi, H., Kamiya, A., Ishiguro, A., Suzuki, A. C., Saitou, N., Toyoda, A. and Aruga, J. (2008). Conservation and diversification of Msx protein in metazoan evolution. *Mol Biol Evol* 25, 69-82.

Takahashi, T., Guron, C., Shetty, S., Matsui, H. and Raghow, R. (1997). A minimal murine Msx-1 gene promoter. Organization of its cis-regulatory motifs and their role in transcriptional activation in cells in culture and in transgenic mice. *J Biol Chem* 272, 22667-78.

Takaoka, K. and Hamada, H. (2012). Cell fate decisions and axis determination in the early mouse embryo. *Development* 139, 3-14.

Takaoka, K., Yamamoto, M. and Hamada, H. (2011). Origin and role of distal visceral endoderm, a group of cells that determines anterior-posterior polarity of the mouse embryo. *Nat Cell Biol* 13, 743-52.

Tam, P. P. and Behringer, R. R. (1997). Mouse gastrulation: the formation of a mammalian body plan. *Mech Dev* 68, 3-25.

Tanabe, Y. and Jessell, T. M. (1996). Diversity and pattern in the developing spinal cord. *Science* 274, 1115-23.

Tanaka, K., Kitagawa, Y. and Kadowaki, T. (2002). Drosophila segment polarity gene product porcupine stimulates the posttranslational N-glycosylation of wingless in the endoplasmic reticulum. *J Biol Chem* 277, 12816-23.

Taneyhill, L. A. and Bronner-Fraser, M. (2005). Dynamic alterations in gene expression after Wnt-mediated induction of avian neural crest. *Mol Biol Cell* 16, 5283-93.

Tao, Q., Yokota, C., Puck, H., Kofron, M., Birsoy, B., Yan, D., Asashima, M., Wylie, C. C., Lin, X. and Heasman, J. (2005). Maternal wnt11 activates the canonical wnt signaling pathway required for axis formation in Xenopus embryos. *Cell* 120, 857-71.

Tassabehji, M., Newton, V. E. and Read, A. P. (1994). Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. *Nat Genet* 8, 251-5.

Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P. and Strachan, T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* 355, 635-6.

Teng, L., Mundell, N. A., Frist, A. Y., Wang, Q. and Labosky, P. A. (2008). Requirement for Foxd3 in the maintenance of neural crest progenitors. *Development* 135, 1615-24.

Theveneau, E. and Mayor, R. (2012). Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration. *Dev Biol* 366, 34-54.

Thibaudeau, G., Holder, S. and Gerard, P. (1998). Anterior/posterior influences on neural crest-derived pigment cell differentiation. *Pigment Cell Res* 11, 189-97.

Thomas, A. J. and Erickson, C. A. (2009). FOXD3 regulates the lineage switch between neural crest-derived glial cells and pigment cells by repressing MITF through a non-canonical mechanism. *Development* 136, 1849-58.

Thomas, P. Q., Brown, A. and Beddington, R. S. (1998). Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. *Development* 125, 85-94.

Thomas, S., Thomas, M., Wincker, P., Babarit, C., Xu, P., Speer, M. C., Munnich, A., Lyonnet, S., Vekemans, M. and Etchevers, H. C. (2008). Human neural crest cells display molecular and phenotypic hallmarks of stem cells. *Hum Mol Genet* 17, 3411-25.

Tompers, D. M., Foreman, R. K., Wang, Q., Kumanova, M. and Labosky, P. A. (2005). Foxd3 is required in the trophoblast progenitor cell lineage of the mouse embryo. *Dev Biol* 285, 126-37.

Tozer, S., Le Dreau, G., Marti, E. and Briscoe, J. (2013). Temporal control of BMP signalling determines neuronal subtype identity in the dorsal neural tube. *Development* 140, 1467-74.

Tribulo, C., Aybar, M. J., Nguyen, V. H., Mullins, M. C. and Mayor, R. (2003). Regulation of Msx genes by a Bmp gradient is essential for neural crest specification. *Development* 130, 6441-52.

Trinh, K. Y., Jin, T. and Drucker, D. J. (1999). Identification of domains mediating transcriptional activation and cytoplasmic export in the caudal homeobox protein Cdx-3. *J Biol Chem* 274, 6011-9.

Urist, M. R. and Strates, B. S. (1971). Bone morphogenetic protein. *J Dent Res* 50, 1392-406.

van Amerongen, R., Mikels, A. and Nusse, R. (2008). Alternative wnt signaling is initiated by distinct receptors. *Sci Signal* 1, re9.

van de Ven, C., Bialecka, M., Neijts, R., Young, T., Rowland, J. E., Stringer, E. J., Van Rooijen, C., Meijlink, F., Novoa, A., Freund, J. N. et al. (2011). Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone. *Development* 138, 3451-62.

van de Wetering, M., Cavallo, R., Dooijes, D., van Beest, M., van Es, J., Loureiro, J., Ypma, A., Hursh, D., Jones, T., Bejsovec, A. et al. (1997). Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene dTCF. *Cell* 88, 789-99.

van den Akker, E. (1999). Targeted inactivation of Hoxb8 affects survival of a spinal ganglion and causes aberrant limb reflexes. *Mech Dev* 89, 103-114.

van den Akker, E., Forlani, S., Chawengsaksophak, K., de Graaff, W., Beck, F., Meyer, B. I. and Deschamps, J. (2002). Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development* 129, 2181-93.

van den Boogaard, M. J., Dorland, M., Beemer, F. A. and van Amstel, H. K. (2000). MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. *Nat Genet* 24, 342-3.

van Nes, J., de Graaff, W., Lebrin, F., Gerhard, M., Beck, F. and Deschamps, J. (2006). The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. *Development* 133, 419-28.

van Rooijen, C., Simmini, S., Bialecka, M., Neijts, R., van de Ven, C., Beck, F. and Deschamps, J. (2012). Evolutionarily conserved requirement of Cdx for post-occipital tissue emergence. *Development* 139, 2576-83.

Vance, K. W. and Goding, C. R. (2004). The transcription network regulating melanocyte development and melanoma. *Pigment Cell Res* 17, 318-25.

Vastardis, H., Karimbux, N., Guthua, S. W., Seidman, J. G. and Seidman, C. E. (1996). A human MSX1 homeodomain missense mutation causes selective tooth agenesis. *Nat Genet* 13, 417-21.

Verzi, M. P., Hatzis, P., Sulahian, R., Philips, J., Schuijers, J., Shin, H., Freed, E., Lynch, J. P., Dang, D. T., Brown, M. et al. (2010a). TCF4 and CDX2, major transcription factors for intestinal function, converge on the same cis-regulatory regions. *Proc Natl Acad Sci U S A* 107, 15157-62.

Verzi, M. P., Shin, H., He, H. H., Sulahian, R., Meyer, C. A., Montgomery, R. K., Fleet, J. C., Brown, M., Liu, X. S. and Shivdasani, R. A. (2010b). Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. *Dev Cell* 19, 713-26.

Verzi, M. P., Shin, H., Ho, L. L., Liu, X. S. and Shivdasani, R. A. (2011). Essential and redundant functions of caudal family proteins in activating adult intestinal genes. *Mol Cell Biol* 31, 2026-39.

Verzi, M. P., Shin, H., San Roman, A. K., Liu, X. S. and Shivdasani, R. A. (2013). Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. *Mol Cell Biol* 33, 281-92.

Waddington, C. H. (1933). Induction by the primitive streak and its derivatives in the chick. *J. Exp. Biol.* 10, 38-46.

Waddington, C. H. (1936). Experiments on determination in the rabbit embryo. *Arch. Biol.* 48, 273-290.

Wallingford, J. B. and Habas, R. (2005). The developmental biology of Dishevelled: an enigmatic protein governing cell fate and cell polarity. *Development* 132, 4421-36.

Wallingford, J. B. and Harland, R. M. (2001). Xenopus Dishevelled signaling regulates both neural and mesodermal convergent extension: parallel forces elongating the body axis. *Development* 128, 2581-92.

Walsh, D. W., Godson, C., Brazil, D. P. and Martin, F. (2010). Extracellular BMP-antagonist regulation in development and disease: tied up in knots. *Trends Cell Biol* 20, 244-56.

Wang, P., Liu, T., Li, Z., Ma, X. and Jin, T. (2006). Redundant and synergistic effect of Cdx-2 and Brn-4 on regulating proglucagon gene expression. *Endocrinology* 147, 1950-8.

Wang, Q., Fang, W. H., Krupinski, J., Kumar, S., Slevin, M. and Kumar, P. (2008a). Pax genes in embryogenesis and oncogenesis. *J Cell Mol Med* 12, 2281-94.

Wang, W., Chen, X., Xu, H. and Lufkin, T. (1996). Msx3: a novel murine homologue of the Drosophila msh homeobox gene restricted to the dorsal embryonic central nervous system. *Mech Dev* 58, 203-15.

Wang, X. D., Morgan, S. C. and Loeken, M. R. (2011). Pax3 Stimulates p53 Ubiquitination and Degradation Independent of Transcription. *PLoS One* 6, e29379.

Wang, Y., Yabuuchi, A., McKinney-Freeman, S., Ducharme, D. M., Ray, M. K., Chawengsaksophak, K., Archer, T. K. and Daley, G. Q. (2008b). Cdx gene deficiency compromises embryonic hematopoiesis in the mouse. *Proc Natl Acad Sci U S A* 105, 7756-61.

Wassarman, K. M., Lewandoski, M., Campbell, K., Joyner, A. L., Rubenstein, J. L., Martinez, S. and Martin, G. R. (1997). Specification of the anterior hindbrain and establishment of a normal mid/hindbrain organizer is dependent on Gbx2 gene function. *Development* 124, 2923-34.

Wehner, P., Shnitsar, I., Urlaub, H. and Borchers, A. (2011). RACK1 is a novel interaction partner of PTK7 that is required for neural tube closure. *Development* 138, 1321-7.

Weiss, M. B., Abel, E. V., Dadpey, N. and Aplin, A. E. (2014a). FOXD3 Modulates Migration Through Direct Transcriptional Repression of TWIST1 in melanoma. *Mol Cancer Res*.

Weiss, M. B., Abel, E. V., Dadpey, N. and Aplin, A. E. (2014b). FOXD3 modulates migration through direct transcriptional repression of TWIST1 in melanoma. *Mol Cancer Res* 12, 1314-23.

Wen, X. Z., Akiyama, Y., Baylin, S. B. and Yuasa, Y. (2006). Frequent epigenetic silencing of the bone morphogenetic protein 2 gene through methylation in gastric carcinomas. *Oncogene* 25, 2666-73.

Wessely, O., Agius, E., Oelgeschlager, M., Pera, E. M. and De Robertis, E. M. (2001). Neural induction in the absence of mesoderm: beta-catenin-dependent expression of secreted BMP antagonists at the blastula stage in Xenopus. *Dev Biol* 234, 161-73.

Weston, J. A. (1991). Sequential segregation and fate of developmentally restricted intermediate cell populations in the neural crest lineage. *Curr Top Dev Biol* 25, 133-53.

Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., Lee, T. I. and Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307-19.

Wiggan, O., Fadel, M. P. and Hamel, P. A. (2002). Pax3 induces cell aggregation and regulates phenotypic mesenchymal-epithelial interconversion. *J Cell Sci* 115, 517-29.

Wilkie, A. O., Tang, Z., Elanko, N., Walsh, S., Twigg, S. R., Hurst, J. A., Wall, S. A., Chrzanowska, K. H. and Maxson, R. E., Jr. (2000). Functional haploinsufficiency of the human homeobox gene MSX2 causes defects in skull ossification. *Nat Genet* 24, 387-90.

Wilkinson, D. G. (1992). In " *In Situ* Hybridisation. A Practical Approach.", (ed. I. Press). New York.

Williams, M., Yen, W., Lu, X. and Sutherland, A. (2014). Distinct apical and basolateral mechanisms drive planar cell polarity-dependent convergent extension of the mouse neural plate. *Dev Cell* 29, 34-46.

Wilson, D. B. (1974). Proliferation in the neural tube of the splotch (Sp) mutant mouse. *J Comp Neurol* 154, 249-55.

Wilson, P. A. and Hemmati-Brivanlou, A. (1995). Induction of epidermis and inhibition of neural fate by Bmp-4. *Nature* 376, 331-3.

Wilson, Y. M., Richards, K. L., Ford-Perriss, M. L., Panthier, J. J. and Murphy, M. (2004). Neural crest cell lineage segregation in the mouse neural tube. *Development* 131, 6153-62.

Witek, M. E., Park, J. and Waldman, S. A: . (2012). CDX2. In Yusuf D et al.: The Transcription Factor Encyclopedia. *Genome Biology , 13* R24.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* 25, 1-47.

Wolpert, L. (1996). One hundred years of positional information. *Trends Genet* 12, 359-64.

Wu, C. I., Hoffman, J. A., Shy, B. R., Ford, E. M., Fuchs, E., Nguyen, H. and Merrill, B. J. (2012). Function of Wnt/beta-catenin in counteracting Tcf3 repression through the Tcf3-beta-catenin interaction. *Development* 139, 2118-29.

Wu, J. and Mlodzik, M. (2009). A quest for the mechanism regulating global planar cell polarity of tissues. *Trends Cell Biol* 19, 295-305.

Wu, M., Li, J., Engleka, K. A., Zhou, B., Lu, M. M., Plotkin, J. B. and Epstein, J. A. (2008). Persistent expression of Pax3 in the neural crest causes cleft palate and defective osteogenesis in mice. *J Clin Invest* 118, 2076-87.

Wuyts, W., Reardon, W., Preis, S., Homfray, T., Rasore-Quartino, A., Christians, H., Willems, P. J. and Van Hul, W. (2000). Identification of mutations in the MSX2 homeobox gene in families affected with foramina parietalia permagna. *Hum Mol Genet* 9, 1251-5.

Xu, F., Li, H. and Jin, T. (1999). Cell type-specific autoregulation of the Caudal-related homeobox gene Cdx-2/3. *J Biol Chem* 274, 34310-6.

Yajima, I., Endo, K., Sato, S., Toyoda, R., Wada, H., Shibahara, S., Numakunai, T., Ikeo, K., Gojobori, T., Goding, C. R. et al. (2003). Cloning and functional analysis of ascidian Mitf in vivo: insights into the origin of vertebrate pigment cells. *Mech Dev* 120, 1489-504.

Yamamoto, M., Saijoh, Y., Perea-Gomez, A., Shawlot, W., Behringer, R. R., Ang, S. L., Hamada, H. and Meno, C. (2004). Nodal antagonists regulate formation of the anteroposterior axis of the mouse embryo. *Nature* 428, 387-92.

Yamanaka, Y., Lanner, F. and Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* 137, 715-24.

Yamanaka, Y., Tamplin, O. J., Beckers, A., Gossler, A. and Rossant, J. (2007). Live imaging and genetic analysis of mouse notochord formation reveals regional morphogenetic mechanisms. *Dev Cell* 13, 884-96.

Yamashita, H., ten Dijke, P., Huylebroeck, D., Sampath, T. K., Andries, M., Smith, J. C., Heldin, C. H. and Miyazono, K. (1995). Osteogenic protein-1 binds to activin type II receptors and induces certain activin-like effects. *J Cell Biol* 130, 217-26.

Yang, Y., Hwang, C. K., Junn, E., Lee, G. and Mouradian, M. M. (2000). ZIC2 and Sp3 repress Sp1-induced activation of the human D1A dopamine receptor gene. *J Biol Chem* 275, 38863-9.

Yang, Y. P. and Klingensmith, J. (2006). Roles of organizer factors and BMP antagonism in mammalian forebrain establishment. *Dev Biol* 296, 458-75.

Yasunami, M., Suzuki, K., Houtani, T., Sugimoto, T. and Ohkubo, H. (1995). Molecular characterization of cDNA encoding a novel protein related to transcriptional enhancer factor-1 from neural precursor cells. *J Biol Chem* 270, 18649-54.

Ybot-Gonzalez, P., Cogram, P., Gerrelli, D. and Copp, A. J. (2002). Sonic hedgehog and the molecular regulation of mouse neural tube closure. *Development* 129, 2507-17.

Ybot-Gonzalez, P., Gaston-Massuet, C., Girdler, G., Klingensmith, J., Arkell, R., Greene, N. D. and Copp, A. J. (2007a). Neural plate morphogenesis during mouse neurulation is regulated by antagonism of Bmp signalling. *Development* 134, 3203-11.

Ybot-Gonzalez, P., Savery, D., Gerrelli, D., Signore, M., Mitchell, C. E., Faux, C. H., Greene, N. D. and Copp, A. J. (2007b). Convergent extension, planar-cell-polarity signalling and initiation of mouse neural tube closure. *Development* 134, 789-99.

Yen, W. W., Williams, M., Periasamy, A., Conaway, M., Burdsal, C., Keller, R., Lu, X. and Sutherland, A. (2009). PTK7 is essential for polarized cell motility and convergent extension during mouse gastrulation. *Development* 136, 2039-48.

Yoshikawa, Y., Fujimori, T., McMahon, A. P. and Takada, S. (1997). Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. *Dev Biol* 183, 234-42.

Young, T., Rowland, J. E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J. N. et al. (2009). Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev Cell* 17, 516-26.

Yu, J. K., Meulemans, D., McKeown, S. J. and Bronner-Fraser, M. (2008). Insights from the amphioxus genome on the origin of vertebrate neural crest. *Genome Res* 18, 1127-32.

Zaret, K. S. and Carroll, J. S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 25, 2227-41.

Zeng, L., Fagotto, F., Zhang, T., Hsu, W., Vasicek, T. J., Perry, W. L., 3rd, Lee, J. J., Tilghman, S. M., Gumbiner, B. M. and Costantini, F. (1997). The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell* 90, 181-92.

Zhao, T., Gan, Q., Stokes, A., Lassiter, R. N., Wang, Y., Chan, J., Han, J. X., Pleasure, D. E., Epstein, J. A. and Zhou, C. J. (2013). beta-catenin regulates Pax3 and Cdx2 for caudal neural tube closure and elongation. *Development*.

Zhao, T., Gan, Q., Stokes, A., Lassiter, R. N., Wang, Y., Chan, J., Han, J. X., Pleasure, D. E., Epstein, J. A. and Zhou, C. J. (2014). beta-catenin regulates Pax3 and Cdx2 for caudal neural tube closure and elongation. *Development* 141, 148-57.

Zhao, X. and Duester, G. (2009). Effect of retinoic acid signaling on Wnt/beta-catenin and FGF signaling during body axis extension. *Gene Expr Patterns* 9, 430-5.

Zhao, Y., Tong, C. and Jiang, J. (2007). Hedgehog regulates smoothened activity by inducing a conformational switch. *Nature* 450, 252-8.

Zimmerman, L. B., De Jesus-Escobar, J. M. and Harland, R. M. (1996). The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell* 86, 599-606.